Virology 483 (2015) 236–252

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

The HIV-1 leader RNA is exquisitely sensitive to structural changes

Nikki van Bel, Anouar Ghabri, Atze T. Das, Ben Berkhout *

Laboratory of Experimental Virology, Department of Medical Microbiology, Center for Infection and Immunity Amsterdam (CINIMA), Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 13 December 2014
Returned to author for revisions
5 March 2015
Accepted 27 March 2015
Available online 15 May 2015

Keywords:
HIV-1 RNA
RNA sequence and structure
Virus evolution
RNA architecture

ABSTRACT

The untranslated leader of the HIV-1 RNA genome is highly structured and contains multiple replication signals. We probed in detail the sequence requirements of a small single-stranded domain using a combination of in silico, in vitro and in vivo virus experiments. Although 'structure-neutral' mutations can be designed by RNA prediction algorithms, experimental follow-up studies nearly always demonstrate local or regional RNA structure changes. Our results indicate that the wild-type HIV-1 RNA sequence has been selected from total sequence space as a unique solution to present critical replication signals in the context of a complex leader structure with small intervening single-stranded segments.

Introduction

The sequence of the 5′ untranslated leader of the human immunodeficiency virus type 1 (HIV-1) RNA genome is highly conserved among virus isolates (Foley et al., 2013). This 368-nucleotide (nt) leader RNA can fold several hairpin structures that regulate multiple processes of the viral replication cycle (Fig. 1A). For instance, the polyA hairpin encompasses the polyadenylation signal that needs to be suppressed at this position, the DIS (Dimerization Initiation Site) hairpin initiates the process of HIV-1 RNA dimerization, the SD (Splice Donor) hairpin includes the major splice donor signal to allow regulation of RNA splicing and the Ψ hairpin is part of the signal for packaging of the HIV-1 genomic RNA in virion particles (Abbink and Berkhout, 2003; Lu et al., 2011b). Dimerization of the HIV-1 genome is thought to be a prerequisite for RNA packaging in virions. The 6-nt palindromic sequence in the apical loop of the DIS hairpin can establish intermolecular base pairs between two RNA genomes, thereby forming a labile kissing-loop (KL) dimer (Haddrick et al., 1996; Laughrea and Jette, 1996; Paillart et al., 2004; Skripkin et al., 1994) (Fig. 1B). This labile dimer is subsequently stabilized in virions to become a mature dimer. In vitro studies indicated that the KL dimer can be converted into an extended dimer (ED) duplex (Huthoff and Berkhout, 2001b; Muriaux et al., 1996a, 1996b) (Fig. 1B), but the role of the ED structure in vivo remains uncertain.

Some of the functions of the HIV-1 leader RNA, like RNA dimerization, are regulated by alternative RNA conformations (Abbink and Berkhout, 2003; Abbink et al., 2005; Huthoff and Berkhout, 2001b, 2002; Ooms et al., 2004). Dimerization is restricted in the ground-state LDI (long-distance interaction) conformation because DIS nucleotides are paired to the upstream polyA hairpin sequences, thus precluding DIS hairpin formation (Fig. 1A) (Ooms et al., 2004). NC protein-mediated induction of the BMH (branched multiple hairpin) conformation exposes the DIS hairpin for KL dimer formation (Fig. 1A). This riboswitch may regulate additional mechanisms of the viral replication cycle and allow coordination of different steps, e.g. RNA dimerization and subsequent packaging (Abbink et al., 2005; Ooms et al., 2004). Recently, an alternative LDI conformation has been proposed in which the DIS is blocked by pairing of the palindromic loop with U5 sequences (Lu et al., 2011a). Also slightly different BMH-like conformations have been suggested (Heng et al., 2012; Lu et al., 2011a; Sakuragi et al., 2012; Wilkinson et al., 2008).

Much of the knowledge on RNA dimerization has been obtained in vitro, whereas the in vivo situation remains much less clear. Deletions and mutations introduced in the leader frequently cause pleiotropic defects in RNA dimerization and RNA packaging (Andersen et al., 2004; Das et al., 2012; Heng et al., 2012; Huthoff and Berkhout, 2001a; Jalalirad et al., 2012; Shen et al., 2001; Song et al., 2008; Vrolijk et al., 2008). The DIS hairpin and its flanking regions seem to be of special importance for these processes (Haddrick et al., 1996; Laughrea and Jette, 1996; Moore and Hu, 2009; Sakuragi et al., 2003; Shen et al., 2001; Skripkin et al., 1994; van Bel et al., 2014b). The DIS hairpin is dispensable for virus replication in certain cell types (Hill et al., 2003; Jones et al., 2008)

* Corresponding author.
E-mail addresses: n.vanbel@amc.uva.nl (N. van Bel), a_ghabri@live.nl (A. Ghabri), a.t.das@amc.uva.nl (A.T. Das), b.berkhout@amc.uva.nl (B. Berkhout).

http://dx.doi.org/10.1016/j.virol.2015.03.050
0042-6822/

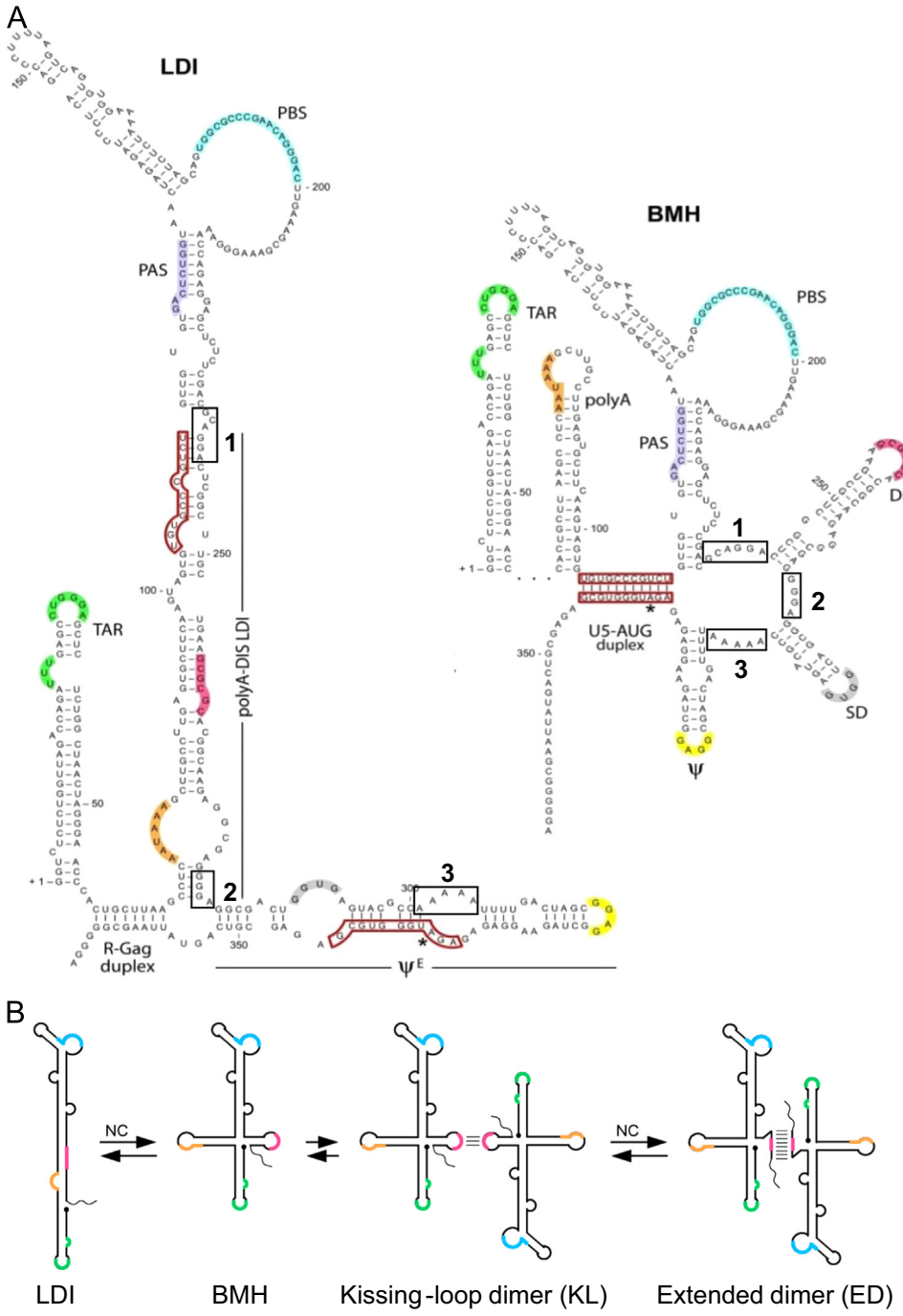

Fig. 1. Secondary structure models and dimerization pathway of the HIV-1 leader RNA. (A) The leader of the HIV-1 RNA genome (nt 1–368) can adopt two conformations: the long-distance interaction (LDI) and the branched multiple hairpin (BMH). Important replication signals are color-coded. The boxes show the positions of the single-stranded regions 1, 2 and 3. Figure adopted from Abbink and Berkhout (2003). (B) Dimerization pathway of HIV-1 RNA. The LDI and BMH conformations are in equilibrium with each other, the LDI conformation is dimerization-incompetent as it does not expose the DIS hairpin. The BMH conformation can dimerize via the exposed DIS hairpin, enabling the formation of a kissing-loop dimer (KL) by basepairing of the palindromic loop sequence of a second HIV-1 genome. The KL dimer can be rearranged and stabilized into an extended dimer (ED) in which intramolecular basepairing in the DIS hairpin is changed into intermolecular basepairing between two DIS hairpins.

and DIS palindrome mutants package a dimeric RNA genome in virions, but with a reduced efficiency (Berkhout and van Wamel, 1996). This has led to the idea that, although the DIS palindrome acts as the primary interaction site, multiple RNA–RNA interaction sites spread along the genome are present in vivo (Fu et al., 1994; Hoglund et al., 1997).

The process of RNA dimerization has been linked to genomic RNA packaging and RNA partner selection, reverse transcription and recombination (Berkhout et al., 1998; Chen et al., 2009; Chin et al., 2005; Moore et al., 2007). For instance, RNA mutants with impaired dimerization capacity often package less of the full-length HIV-1 RNA genome, which is compensated by increased packaging of spliced HIV-1 RNAs (Clever and Parslow, 1997; Das et al., 2012; Didierlaurent et al., 2011; Houzet et al., 2007). Formation of the RNA dimer structure is critical for packaging and RNA molecules engineered with an intramolecular KL interaction are efficiently packaged as monomers (Nikolaitchik et al., 2013; Sakuragi et al., 2001). The labile KL RNA dimer is recognized by the NC domain of the Gag polyprotein (Abd El-Wahab et al., 2014; Jouvenet et al., 2009). Several high affinity binding sites are

present in the leader, including a short sequence motif immediately upstream of the DIS hairpin (Amarasinghe et al., 2000; Darlix et al., 2014; Levin et al., 2005; Wilkinson et al., 2008). A recent study suggested that the lower part of the DIS structure forms the primary binding site for the Gag precursor protein to initiate RNA packaging (Abd El-Wahab et al., 2014). This configuration directly juxtaposes the viral RNA dimerization and packaging signals, which may allow coordination of these processes by RNA structure changes. The RNA–Gag complex traffics to the viral budding site at the cellular membrane where an immature virus particle is assembled (Jouvenet et al., 2009). RNA dimer stabilization occurs during virus maturation and both processes require protease-mediated cleavage of the Gag and Gag–Pol polyproteins (Fu et al., 1994; Ohishi et al., 2011; Song et al., 2007). The exact dimer interactions and RNA transitions that occur during virion maturation remain largely unknown.

Many studies addressed the specific role of the hairpin structures in the HIV-1 leader RNA. In contrast, the highly conserved single-stranded regions flanking the DIS, SD and Ψ hairpins have only been investigated to a limited extent. We recently analyzed the sequence space that is compatible with HIV-1 replication in region 1 (^{237}GCAGGA^{242}, upstream of the DIS hairpin), region 2 (^{278}GGGA^{281}, connecting the DIS and the SD hairpins) and region 3 (^{301}AAAAA^{305}, connecting the SD and the Ψ hairpins) (Fig. 1A) (van Bel et al., 2014a). Replication-competent viruses were selected from an HIV-1 DNA library with randomized segments, which revealed a particularly high selection pressure for region 1. In this study, an in-depth functional analysis of the sequence and structure requirements in regions 1, 2 and 3 was performed using mutants that were randomly picked from these libraries. For replication-impaired mutants, we subsequently selected revertant viruses in long-term cultures. The variants were screened and experimentally tested for the impact on folding of the DIS hairpin and the complete leader RNA. We also performed several functional RNA tests in vitro (RNA structure probing and dimerization) and in vivo (RNA dimerization and packaging in virions) towards a better understanding of these RNA motifs.

Results

HIV-1 region 1/2/3 mutants

To study the role of regions 1, 2 and 3 in the HIV-1 leader RNA, we analyzed a large set of virus mutants (Table 1). These regions are single-stranded in the BMH conformation, but involved in some base pairing interactions in other RNA structure models and isolates (Abbink and Berkhout, 2003; Heng et al., 2012; Lu et al., 2011a, 2011b; Sakuragi et al., 2012; van Bel et al., 2014a; Wilkinson et al., 2008). These mutants were randomly picked from previously created SELEX libraries based on the full-length HIV-1 LAI genome with small segments of randomized sequence (van Bel et al., 2014a). The complete leader RNA was sequenced to confirm the absence of additional mutations outside the randomized stretch. We tested virus production in 293T human embryonal kidney cells and virus replication in SupT1 human T cells. DNA-transfected 293T cells allow high virus production, but no virus replication because of the absence of the CD4 receptor for viral entry. 293T cells were transfected with in total 66 mutant LAI-based virus clones: 22 × region 1 (mutants 1A–1V; Table 1), 23 × region 2 (2A–2W) and 21 × region 3 (3A–3U). Virus production was quantitated by measuring the CA-p24 level in the culture supernatant after 2 days. SupT1 cells do express the CD4 receptor and are susceptible to HIV-1 replication. The replication capacity of the virus mutants was analyzed by infecting SupT1 T cells with equal amounts of 293T-produced virus (based on CA-p24) and viral spread was measured by monitoring the CA-p24 level in the supernatant. For region 1 mutants, no significant virus production defects were apparent (Fig. 2A), but virus replication was frequently hampered. We classified the mutants in three groups based on their replication capacity: no or a minor problem (Fig. 2B), a moderate problem (Fig. 2C) or a severe replication defect (Fig. 2D). Most of the region 2 and 3 mutants showed no significant defect in virus production and virus replication (Fig. 2E, F and H, respectively). Two region 2 mutants (2L and 2T) and one region 3 mutant (3E) formed an exception. No explanation is apparent for the region 3 mutant. Both region 2 mutants introduce an AUG translational start codon upstream of the Gag open reading frame in a favorable Kozak sequence context, which likely inhibits Gag protein translation, virus production and replication (Das et al., 1998; Hache et al., 2009; Kozak, 1991). Several other mutants encode an upstream AUG, apparently without a severe effect on virus production and replication (Table 1). In some of them, a sub-optimal Kozak sequence context may provide a rationale for this.

Table 1
Sequences of region 1, 2 and 3 mutants.

Region 1 mutants [a]	Position (237–242)[b]	Region 2 mutants [a]	Position (278–281)[b]	Region 3 mutants [a]	Position (301–305)[b]
wt	GCAGGA	wt	GGGA	wt	AAAAA
1A	UUUGUA	2A	UGGA	3A	UGGAG
1B	CUCUUC	2B	AAGA	3B	UGAGG
1C	UAAUUG	2C	GAUA	3C	GCGGC
1D	AGUAAU	2D	AUUG	3D	UGGGA
1E	CUUUGA	2E	GUAG	3E	CACGA
1F	AACCCU	2F	UUUC	3F[c]	UU*AUG*
1G[c]	C*AUG*CU	2G	AACA	3G	ACGUG
1H[c]	C*AUG*UA	2H	GGGU	3H	GGACU
1I	AACGAA	2I	UUAA	3I	CGGGA
1J	UAUCGA	2J	UUUG	3J	AGCUU
1K	CUUGAC	2K	UGGG	3K	AUUCG
1L	CAAAGA	2L[c]	G*AUG*	3L	UCGGA
1M	AGCAAC	2M[c]	G*AAU*	3M	AAGGG
1N	AAUACC	2N	UAUC	3N[c]	AC*AUG*
1O	AAGGGA	2O	GAUA	3O	AUAAG
1P	UCACUC	2P	GAGU	3P	GGUAG
1Q	UUCCAC	2Q	GGUA	3Q	UGCAU
1R	CUUCCC	2R	AUAG	3R	GGCUU
1S	ACAGUA	2S	GCCG	3S	GCAUU
1T	AGAUAC	2T[c]	*AUG*A	3T	UUUGA
1U	CCCCCA	2U	UUAG	3U	GAGCG
1V	AACAAU	2V	CAGA		
		2W	UAAA		

[a] Mutant name.
[b] nt differing from wt are underlined for clarity.
[c] Contains potential AUG startcodon (marked in italics).

RNA structure effects

We used RNAstructure software to predict the impact of the region 1, 2 and 3 mutations on the local RNA secondary structure (Reuter and Mathews, 2010). We first tested whether the mutations change the folding of the upstream or downstream positioned hairpin by analyzing a small segment consisting of an individual hairpin (DIS, SD or Ψ) with the flanking single-stranded regions. The thermodynamic stability (Δ*G*) of the most stable predicted structure is listed (Tables 2–4). The ΔG_{DIS} values of the wt hairpins are as follow: DIS, −12.9; SD, −9.6 and Ψ, −7.6 kcal/mol. Strikingly, the majority of the region 1 mutations extend and stabilize the DIS hairpin via the formation of additional base pairs with region 2 (11 mutants increase the DIS hairpin stability with >25%, i.e. $\Delta G_{DIS} < -16.2$ kcal/mol; Table 2). For region 2, only a few mutations stabilize the DIS hairpin to some extent (no mutants with a $\Delta G_{DIS} < -16.2$) and no or relatively

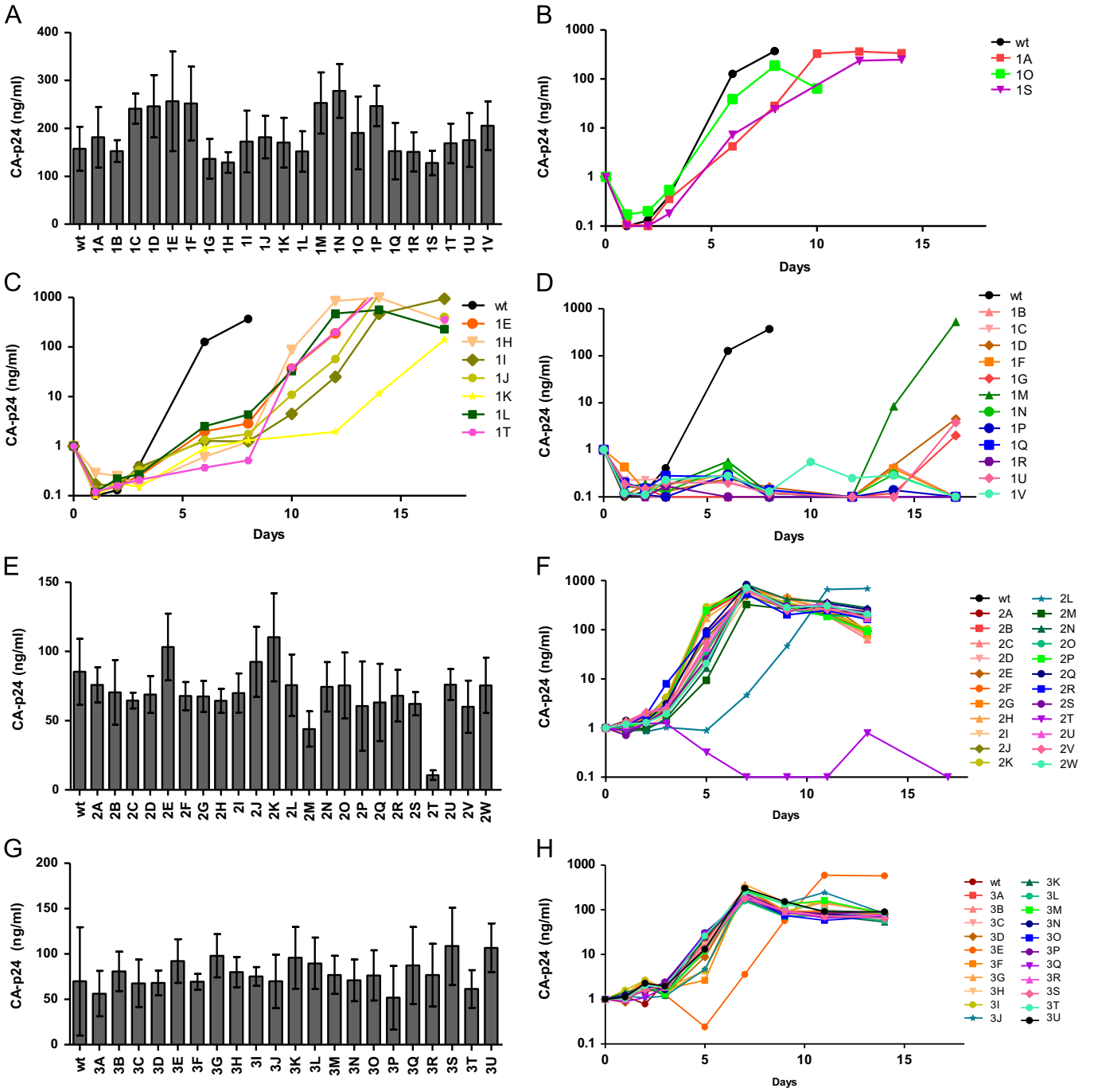

Fig. 2. Virus production and replication capacity of region 1, 2 and 3 HIV-1 mutants. Virus production was measured upon transfection of 293T cells with plasmid DNA encoding region 1 (A), region 2 (E) or region 3 (G) mutants, using the calcium-phosphate method. After 48 h the CA-p24 level in the supernatant was measured. For virus replication, SupT1 T cells were infected with equivalent amounts of 293T-produced virus. Shown is the mean ($N=3$–5) with SD. Virus spread was assessed by measuring the CA-p24 level in the culture supernatant. The region 1 mutants are shown in three graphs, depending on their replication capacity: no or minimal defects (B; CA-p24 level at day 7 > 50% of the wt level); moderate defects (C; CA-p24 level at day 7 < 50% and at day 10 > 35% of the maximal wt level); strong defects or replication incompetent (D; undetectable CA-p24 levels or CA-p24 detectable after day 10). (F–H) Virus replication of region 2 mutants (F) and region 3 mutants (H). Replication of all virus mutants was assayed at least two times, and similar replication curves were obtained in independent experiments.

small effects were scored on the SD hairpin stability (Table 3). Region 3 mutations did not or modestly affect the SD and Ψ hairpin stability (Table 4).

We next tested the effect on the structure of the complete leader RNA (nt 1–368). The ΔG value was calculated for the wt and mutant sequences in both the LDI (ΔG_{LDI}) and BMH (ΔG_{BMH}) conformations. The $\Delta\Delta G_{LDI-BMH}$ value describes the equilibrium between these two conformations. A $\Delta\Delta G_{LDI-BMH}$ value of −5.6 kcal/mol for wt indicates that HIV-1 LAI RNA preferably folds the dimerization-incompetent LDI conformation, but this may vary for other virus isolates (Kasprzak et al., 2005). A positive $\Delta\Delta G_{LDI-BMH}$ value ($\Delta G_{BMH} < \Delta G_{LDI}$) indicates a shift towards preferential folding of the dimerization-competent BMH conformation. We calculated the $\Delta\Delta G_{LDI-BMH}$ value for all region 1, 2 and 3 mutants (Tables 2–4). Nearly half of the region 1 mutants preferentially fold the BMH conformation ($\Delta\Delta G_{LDI-BMH} > 0$), and all these mutants demonstrated DIS hairpin stabilization (Table 2). Such a strong LDI-to-BMH shift is predicted for only two region

Table 2
RNA secondary structure and virus replication of region 1 mutants.

Mutant	Sequence[a]	Replication[b]	ΔG_{DIS}[c]	ΔG_{LDI}[c]	ΔG_{BMH}[c]	$\Delta\Delta G_{LDI-BMH}$[d]
wt	GCAGGA	++	−12.9	−139.8	−134.2	−5.6
1O	AAGGGA	++	−12.9	−142.3	−136.4	−5.9
1A	UUUGUA	+	−12.9	−139.2	−135.2	−4.0
1S	ACAGUA	+	−12.9	−139.2	−135.6	−3.6
1T	AGAUAC	+/−	−16.2	−141.1	−136.7	−4.4
1H	CAUGUA	+/−	−12.9	−139.2	−135.5	−3.7
1I	AACGAA	+/−	−12.9	−138.8	−135.9	−2.9
1L	CAAAGA	+/−	−12.9	−138.3	−135.7	−2.6
1J	UAUCGA	+/−	−14.4	−136.1	−136.8	0.7
1K	CUUGAC	+/−	−16.2	−137.6	No BMH	No BMH
1E	CUUUGA	+/−	−12.9	−136.1	No BMH	No BMH
1V	AACAAU	−	−13.1	−138.3	−133.8	−4.5
1M	AGCAAC	−	−16.4	−138.3	−140.2	1.9
1D	AGUAAU	−−	−13.1	−138.3	−134.3	−4.0
1B	CUCUUC	−−	−17.1	−135.8	−133.6	−2.2
1C	UAAUUG	−−	−13.5	−136.3	−137.0	0.7
1G	CAUGCU	−−	−16.3	−136.7	−139.9	3.2
1N	AAUACC	−−	−19.5	−136.4	−141.6	5.2
1F	AACCCU	−−	−19.7	−136.6	−142.3	5.7
1P	UCACUC	−−	−19.3	−136.2	−142.5	6.3
1Q	UUCCAC	−−	−19.5	−137.4	−144.5	7.1
1U	CCCCCA	−−	−22.7	−134.1	−147.9	13.8
1R	CUUCCC	−−	−23.3	No LDI	−147.6	No LDI

The mutants were sorted first on replication capacity and within groups sorted on $\Delta\Delta G_{LDI-BMH}$.

[a] Nucleotides differing from wt underlined.

[b] Virus replication scored by ++ (day 6/7: CA-p24 > 95% of maximum of wt), + (day 6: CA-p24 > 50% of maximum of wt), +/− (day 10: CA-p24 > 35% of maximum of wt), − (detectable CA-p24 level after day 10), and −− (no CA-p24 detected).

[c] ΔG in kcal/mol of the DIS hairpin and the complete leader in LDI and BMH conformation as predicted by RNAstructure.

[d] $\Delta\Delta G$ in kcal/mol of the LDI–BMH equilibrium.

Table 3
RNA secondary structure and virus replication of region 2 mutants.

Mutant	Sequence[a]	Replication[b]	ΔG_{DIS}[c]	ΔG_{SD}[c]	ΔG_{LDI}[c]	ΔG_{BMH}[c]	$\Delta\Delta G_{LDI-BMH}$[d]
wt	GGGA	++	−12.9	−9.6	−139.8	−134.2	−5.6
2A[e]	UGGA	++	−14.2	−9.6	−139.8	−135.2	−4.6
2B	AAGA	++	−12.4	−9.6	−139.0	−134.5	−4.5
2C	GAUA	++	−12.9	−9.6	−139.1	−135.0	−4.1
2V	CAGA	+	−12.4	−9.6	−139.0	−128.6	−10.4
2E	GUAG	+	−12.9	−9.6	−140.0	−135.2	−4.8
2G	AACA	+	−12.4	−9.6	−139.3	−134.5	−4.8
2P	GAGU	+	−12.9	−10.3	−139.8	−135.2	−4.6
2M	GAAU	+	−12.9	−9.6	−139.8	−135.7	−4.1
2O	GAUA	+	−12.9	−9.6	−139.1	−135.0	−4.1
2A	GGUA	+	−12.9	−9.6	−139.2	−135.2	−4.0
2H	GGGU	+	−12.9	−10.3	−140.1	−136.3	−3.8
2S	GCCG	+	−15.8	−9.6	−141.6	−138.1	−3.5
2D	AUUG	+	−12.4	−9.6	−139.0	−136.7	−2.3
2K	UGGG	+	−14.2	−9.6	−138.9	−136.6	−2.3
2R	AUAG	+	−12.4	−9.6	−138.0	−135.7	−2.3
2N	UAUC	+	−14.1	−9.6	−138.9	−136.7	−2.2
2W	UAAA	+	−14.4	−9.6	−138.0	−136.4	−1.6
2I	UUAA	+	−15.4	−9.6	−138.0	−137.8	−0.2
2U	UUAG	+	−15.4	−9.6	−138.0	−137.8	−0.2
2F	UUUC	+	−15.5	−9.6	−138.0	−138.9	0.9
2J	UUUG	+	−15.5	−9.6	−138.0	−138.8	0.8
2T	AUGA	−−	−12.4	−9.6	−140.5	−135.9	−4.6
2L	GAUG	−−	−12.9	−10.3	−139.1	−136.1	−3.0

[a] Nucleotides differing from wt underlined.

[b] Virus replication scored by ++ (day 6/7: CA-p24 > 95% of maximum of wt), + (day 6: CA-p24 > 50% of maximum of wt), +/− (day 10: CA-p24 > 35% of maximum of wt), −− (no CA-p24 detected).

[c] ΔG in kcal/mol of the DIS and SD hairpins and the complete leader in LDI and BMH conformation as predicted by RNAstructure.

[d] $\Delta\Delta G$ in kcal/mol of the LDI–BMH equilibrium.

[e] Shows a rearranged base-pairing scheme of the DIS hairpin in the RNAstructure prediction. The nucleotides which appear single-stranded in the DIS hairpin of the wt (positions 247, 255–263, and 271–273) were forced to be single stranded in the predictions.

Table 4
RNA secondary structure and virus replication of region 3 mutants.

Mutant	Sequence[a]	Replication[b]	ΔG_{SD}[c]	ΔG_{Ψ}[c]	ΔG_{LDI}[c]	ΔG_{BMH}[c]	$\Delta\Delta G_{LDI-BMH}$[d]
wt	AAAAA	++	−9.6	−7.6	−139.8	−134.2	−5.6
3P	GGUAG	++	−9.6	−7.6	−142.1	−134.3	−7.8
3F	UUAUG	++	−11.9	−7.6	−142.3	−135.6	−6.7
3J	AGCUU	++	−9.9	−10.4	−142.6	−135.9	−6.7
3G	ACGUG	++	−9.6	−7.6	−143.9	−137.3	−6.6
3L	UCGGA	++	−14.1	−7.6	−144.3	−138.5	−5.8
3S	GCAUU	++	−9.6	−9.1	−144.3	−139.3	−5.0
3I	CGGGA	++	−8.9	−7.6	−146.0	−141.2	−4.8
3K	AUUCG	++	−11.0	−7.6	−141.9	−137.3	−4.6
3Q	UGCAU	++	−10.7	−8.9	−140.8	−136.4	−4.4
3R	GGCUU	++	−9.9	−10.4	−144.3	−139.3	−4.0
3C	GCGGC	++	−9.6	−9.6	−142.0	−137.8	−4.2
3A	UGGAG	++	−10.7	−7.6	−139.7	−136.4	−3.3
3D	UGGGA	++	−10.7	−7.6	−139.7	−136.4	−3.3
3M	AAGGG	++	−9.6	−7.6	−142.5	−139.7	−2.8
3B	UGAGG	++	−10.7	−7.6	−138.9	−136.4	−2.5
3N	ACAUG	++	−9.6	−7.6	−139.9	−137.8	−2.1
3U	GAGCG	++	−9.6	−7.6	−140.7	−138.9	−1.8
3T	UUUGA	++	−12.1	−7.6	−140.7	−138.9	−1.8
3O	AUAAG	++	−9.6	−7.6	−138.3	−137.4	−0.9
3H	GGACU	++	−9.6	−11.2	−137.2	−137.0	−0.2
3E	CACGA	+/−	−8.9	−7.6	−139.2	−138.2	−1.0

[a] Nucleotides differing from wt underlined.

[b] Virus replication scored by ++ (day 6/7: CA-p24 > 95% of maximum of wt), +/− (day 10: CA-p24 > 35% of maximum of wt).

[c] ΔG in kcal/mol of the SD and Ψ hairpins and the complete leader in LDI and BMH conformation as predicted by RNAstructure.

[d] $\Delta\Delta G$ in kcal/mole of the LDI-BMH equilibrium.

2 mutants (Table 3) and not a single region 3 mutant (Table 4). Comparison of the virus replication capacity with the $\Delta\Delta G_{LDI-BMH}$ and ΔG_{DIS} values reveals a possible trend as most of the mutants with a strongly stabilized DIS hairpin and/or with a positive $\Delta\Delta G_{LDI-BMH}$ are replication impaired, especially region 1 mutants. For example, mutants 1P, 1Q and 1U with a positive $\Delta\Delta G_{LDI-BMH}$ value and a stabilized DIS hairpin do not replicate (Table 2). This trend confirms earlier studies showing that stabilization of the DIS hairpin affects the $\Delta\Delta G_{LDI-BMH}$ and inhibits viral replication (Ooms et al., 2004; van Bel et al., 2014a, 2014b). However, replication of some region 1 mutants (1D and 1V) is completely impaired, although ΔG_{DIS} and $\Delta\Delta G_{LDI-BMH}$ are altered only modestly, suggesting that sequence and structural features may be important.

Design of structure-neutral region 1 mutants

To describe sequence-specific effects in more detail, we decided to focus on region 1 because those mutants exhibit the largest replication impact. The strategy was to design additional mutants in which the first or second triplet of region 1, or all 6 nt was substituted without an effect on the predicted RNA structure. RNAstructure analysis demonstrated that most hypothetical triplet mutations affect the ΔG_{DIS} and $\Delta\Delta G_{LDI-BMH}$ values, in agreement with the data shown in Table 2. The screen yielded two structure-neutral triplet mutants and the corresponding double mutant. The wt sequence GCAGGA was mutated into GCAAAG (NB1, mutated nt underlined), CUGGGA (NB2) and – upon combination – CUGAAG (NB3). The predicted DIS hairpin stability (ΔG_{DIS} = −12.7, −12.9 and

−12.7, respectively) and the $\Delta\Delta G_{LDI-BMH}$ of the complete leader ($\Delta\Delta G_{LDI-BMH} = -4.1$, −5.1 and −5.1, respectively) are comparable to wt ($\Delta G_{DIS} = -12.9$ and $\Delta\Delta G_{LDI-BMH} = -5.6$; Table 5).

Table 5
RNA secondary structure and virus replication of designed sequence-only mutants.

Mutant	Sequence[a]	Replication[b]	ΔG_{DIS}[c]	ΔG_{LDI}[c]	ΔG_{BMH}[c]	$\Delta\Delta G_{LDI-BMH}$[d]
wt	GCAGGA	++	−12.9	−139.8	−134.2	−5.6
NB1	GCAAAG	+	−12.7	−138.8	−134.7	−4.1
NB2	CUGGGA	+/−	−12.9	−140.7	−135.6	−5.1
NB3	CUGAAG	−	−12.7	−136.3	−137.3	−5.1

[a] Nucleotides differing from wt underlined.
[b] Virus replication scored by ++ (day 6/7: CA-p24 > 95% of maximum of wt), + (day 6: CA-p24 > 50% of maximum of wt), +/− (day 10: CA-p24 > 35% of maximum of wt), − (detectable CA-p24 level after day 10).
[c] ΔG in kcal/mol of the DIS hairpin and the complete leader in LDI and BMH conformation as predicted by RNAstructure.
[d] $\Delta\Delta G$ in kcal/mol of the LDI–BMH equilibrium.

Virus production of the NB mutants in transfected 293T cells was wt-like (Fig. 3A), but replication on SupT1 T cells was affected with the following ranking order: wt > NB1 > NB2 > NB3 (Fig. 3B). These combined results do suggest a sequence-specific contribution of region 1 to HIV-1 replication. We therefore set out to test these mutants in more detail in vitro (RNA conformation, RNA dimerization) and in vivo (RNA dimerization/packaging in virions and virus evolution).

In vitro RNA dimerization studies with region 1 mutants

RNA dimerization was assessed in vitro using HIV-1 leader RNA transcripts, which were denatured and allowed to refold and dimerize before analysis on an EDTA-containing TBE or Mg^{2+}-containing TBM gel. The TBE gel distinguishes between the BMH and LDI monomer conformations. Furthermore, because EDTA destabilizes KL dimers, only ED are visible (Fig. 3C, left panel) (Berkhout and van Wamel, 2000; Ooms et al., 2004; Vrolijk et al., 2008). Mg^{2+} in the TBM gel stabilizes KL dimers and thus allows

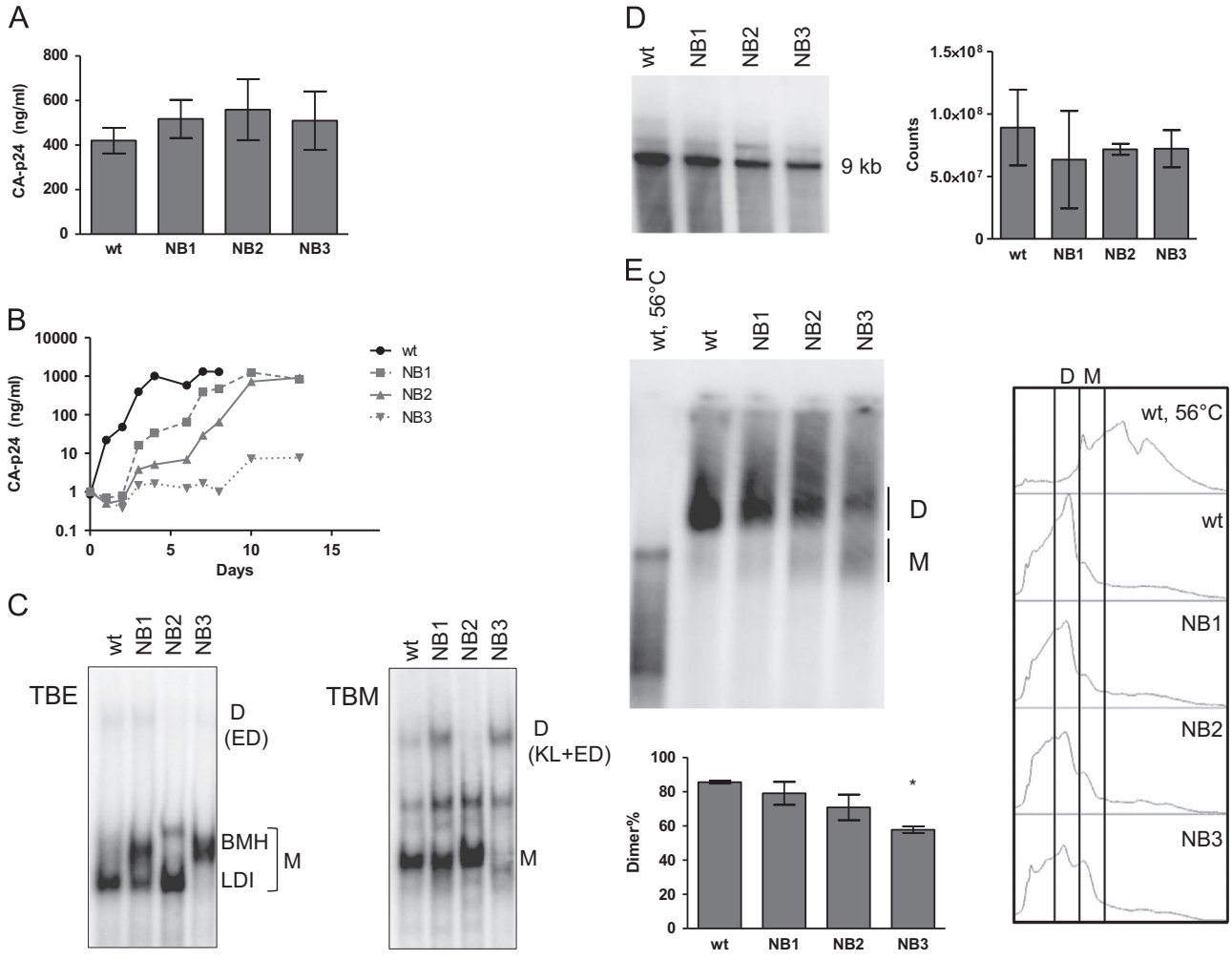

Fig. 3. Characterization of region 1 mutants. (A)Virus production was measured upon transfection of 293T cells with the indicated plasmids. We used lipofectamine2000 for transfection, which yields higher CA-p24 levels than the calcium-phosphate method used in Fig. 2. After 48 h the CA-p24 level in the supernatant was measured. Shown is the average ($N=2$) with SD. (B) Virus replication capacity was analyzed by infection of SupT1 T cells with similar amounts of virus (equivalent to 1 ng CA-p24). Viral spread was monitored by the CA-p24 level in the culture supernatant. Shown is a representative of 3 independent experiments. (C) Leader RNA transcripts (nt 1–368, 200 ng) were heat-denatured and allowed to renature and dimerize before the sample was split and loaded on a TBE or TBM polyacrylamide gel. The positions of the LDI and BMH conformers, the ED, the total dimer (D) and monomer (M) bands are indicated. Shown are representative gels of 4 independent experiments. (D) Packaging of genomic RNA was analyzed by denaturing northern blotting. Viral RNA isolated from equal amounts of virions (equivalent to 250 ng CA-p24) was denatured at 65 °C before loading on a denaturing gel. The position of the full-length genome of 9 kb is indicated. Quantification was performed using ImageQuant. Shown is the average ($N=2$) with SD. (E) Viral RNA dimerization of the wt and mutant viruses was analyzed by non-denaturing northern blotting. Virus-extracted RNA was loaded on a non-denaturing gel omitting the heating step. The position of dimeric RNA (D) and monomeric RNA (M) is indicated. Quantification of the dimer and monomer bands was performed using ImageQuant, shown is the average ($N=2$) with SD. * $p < 0.05$ with Student's t-test. The histogram scans were made using ImageJ.

one to score the total dimer level (Fig. 3C, right panel), but does not distinguish between the BMH and LDI monomers. As previously described (Berkhout and van Wamel, 2000; Huthoff and Berkhout, 2001b), the wt transcript folds the LDI monomer with minute amounts of BMH (TBE gel) and dimers (TBE and TBM gels). The NB1 mutant partially shifts the equilibrium towards the BMH configuration (TBE gel), which exposes the DIS hairpin and causes an increase in KL dimer, but not ED dimer levels. Accordingly, the total KL+ED dimer signal increases on the TBM gel, but the ED level on the TBE gel does not increase. Like the wt, NB2 preferably folds the LDI conformation, but the small BMH fraction migrates slower on the TBE gel than the wt BMH conformer, suggesting an unanticipated conformational effect. The lack of NB2 dimer formation on both gels suggests that this structural effect on the BMH conformer affects its dimerization capacity. The NB3 mutations switch the HIV-1 RNA monomer into the BMH conformation (TBE gel). Consequently, increased KL dimerization is scored (TBM gel), although slightly lower ED levels are observed compared to wt (TBE gel). These results demonstrate that, despite the RNAstructure predictions, the NB region 1 mutations do affect the in vitro leader RNA structure and RNA dimerization.

RNA dimer packaging in virions

We subsequently analyzed packaging and the dimeric state of the HIV-1 RNA inside virus particles. Virus produced upon transfection of 293T cells with wt and mutant HIV-1 plasmids was purified by ultracentrifugation. The viral RNA was gently extracted from virus particles and analyzed on a denaturing gel to quantify the RNA content (Fig. 3D) and RNA dimers were scored on a non-denaturing gel (Fig. 3E). The RNA content of wt and mutant particles was similar, indicating that packaging was not affected (Fig. 3D). Dimerization of the mutant RNA genomes was affected (Fig. 3E). Whereas 83% of the RNA in wt virions was present in the mature, stable dimer conformation, the mutants exhibited a reduced dimer and increased monomer level (Fig. 3E). Notably, NB3 had the lowest dimer content and NB1 the highest, consistent with the ranking order of virus replication.

Probing of the RNA secondary structure

The NB mutants were designed to be neutral to the RNA structure, but we nevertheless scored RNA structure effects in vitro. We therefore probed the RNA secondary structure with the SHAPE method and nucleotide reactivities were implemented in RNAstructure to steer structure prediction. The wt and NB1 RNA showed a similar reactivity profile, which confirms that NB1 does not cause major RNA structural changes, with only minimal effects on DIS hairpin and leader RNA stability (Fig. 4 and Supplementary Table 1). NB3 caused a change in reactivity in the U5-AUG and PBS regions, which indicates alternative base pairing, but no significant effect on the isolated DIS hairpin was observed. In contrast, multiple unexpected structural changes were apparent for NB2. SHAPE-directed RNA structure prediction indicates that NB2 extends the DIS hairpin with three base pairs (^{236}CCU^{238} pairs with ^{278}GGG^{280}), thus creating a new bulge (^{239}GGGA^{242}). The NB2 mutation also altered the reactivity of several nucleotides in the DIS stem (^{249}U and ^{275}GA^{276} have increased reactivity) and loop region. The palindrome ^{257}GCGCGC^{262} in the wt loop is not reactive because of KL formation between two BMH conformers. But ^{259}G in NB2 is more exposed, whereas the adjacent ^{256}A is less reactive (Fig. 4 and Supplementary Table 1). This may suggest that the DIS hairpin is not uniformly present in the BMH conformation, which may explain the slower migrating band on native gels and the loss of dimerization capacity (Fig. 3C). Summarizing, RNA structure probing confirms that the NB mutations do affect the leader RNA structure. Although the effects of NB1 and NB3 were small, they may explain the conformational changes observed on the gels and differences in dimerization observed in vitro and in vivo.

Virus evolution

We performed virus evolution studies with all region 1 mutants to select interesting revertants that could provide further mechanistic insights. We have repeatedly used this powerful genetic method to study the role of RNA elements in HIV-1 replication (Berkhout and Das, 2009). Some HIV-1 mutants replicated slowly initially, but accelerated over time. Others were initially replication-incompetent, but subsequently started spreading. We continuously passaged cell-free virus onto fresh SupT1 cultures and took cell samples for inspection of the leader and sequences encoding the N-terminus of the Gag protein (up to amino acid 15) because these sequences may base pair with leader domains (Abbink and Berkhout, 2003; Sakuragi et al., 2012). Cellular DNA was extracted and the relevant parts of the integrated HIV-1 proviruses were sequenced.

The acquired mutations are listed in Table 6. Both first and second site mutations were observed. First site changes that repair the wt sequence GCAGGA were observed at position 1 (mutant 1V: AACAAU to ***G***ACAAU, mutant nt underlined, revertant nt in bold and italics), position 2 (1V: to A***C***CAAU), position 5 (mutant 1N: AAUACC to AAUA***G***C) and position 6 (mutant 1C: UAAUUG to UAAUU***A***), but not at positions 3 and 4. These results correlate with the results of a SELEX experiment that revealed significant pressure to maintain the wt sequence, especially at positions 1, 2 and 6 (van Bel et al., 2014a). Several first site reversions do not repair the wt sequence. Notably, mutant position 5 was converted to the non-wt U on 3 occasions: once for mutant 1M (AGCAAC to AGCA***U***C) and twice for mutant 1N (AAUACC to AAUA***U***C), consistent with the previous SELEX experiment (van Bel et al., 2014a).

Second site reversions outside region 1 were also frequently observed. Remarkably, we frequently detected changes at ^{106}G and in the region upstream of the PBS. Furthermore, several viruses acquired mutations that caused an amino acid substitution in the N-terminus of the Gag-encoded Matrix protein (mutants 1A, 1C, 1G, 1K, 1L, 1P, 1Q, 1R, 1S, 1W, Table 6). Many of the acquired mutations were recloned in the original mutant constructs, but only a few, mostly the first-site revertants, did significantly improve virus replication (Table 6, last column).

Some viruses started replicating without compensatory mutations in the leader. For these cultures we performed a more extensive sequence analysis of the complete gag gene, because compensatory mutations have previously been observed in other gag regions of leader mutated HIV-1 variants (Liang et al., 1999a, 1998, 1999b). However, this did not reveal candidate compensatory changes.

Revertant analysis of region 1 structure mutants

In earlier studies we noticed that mutants with a stabilized DIS hairpin demonstrated RNA dimerization problems (Abbink et al., 2005; van Bel et al., 2014b). To confirm and extend these findings for a larger mutant set, we selected two mutant-revertant sets for further analysis. First, we analyzed two sets for which RNAstructure predicts that the region 1 mutation causes DIS hairpin stabilization and a shift toward BMH (positive $\Delta\Delta G_{LDI-BMH}$, Table 7), whereas the reversions partially restore the wt characteristics: mutant 1N (AAUACC) with revertants 1N7 (AAUA***U***C) and 1N22 (AAUA***G***C), and mutant 1P (UCACUC) with revertant 1P22 (UCA***A***UC). The 1N and 1P mutations profoundly stabilized the DIS hairpin (more negative ΔG_{DIS}; Table 7) by adding two base pairs (^{241}CC^{242} interacts with

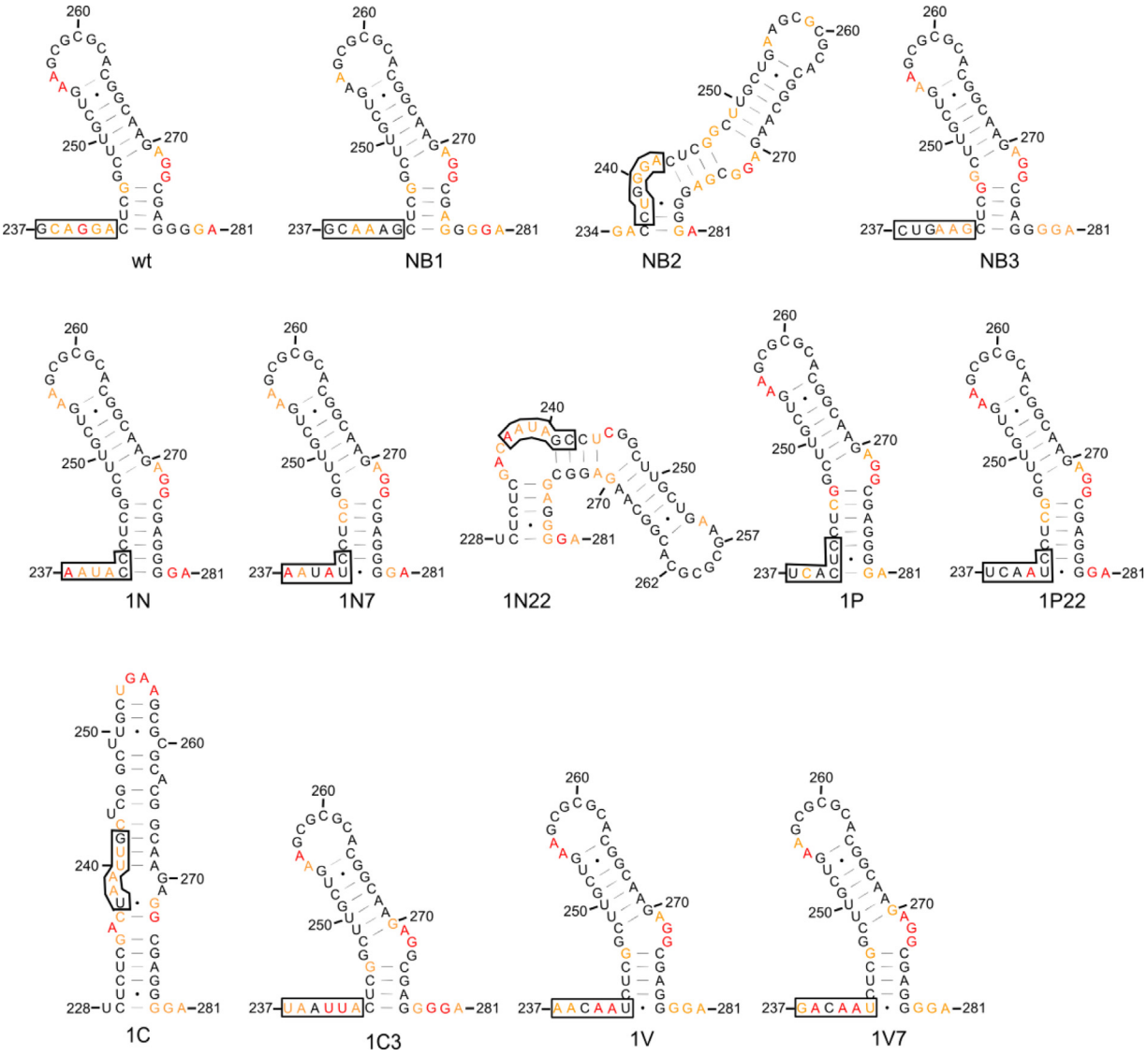

Fig. 4. SHAPE-directed RNA secondary structure prediction of region 1 mutants. Leader RNA transcripts (nt 1–368) were folded in a buffer that forces the RNA into the BMH conformation and consequently induces KL dimerization. SHAPE reactivity data were obtained and used as pseudo-energy constraints in the RNAstructure program to steer the secondary structure prediction. The DIS hairpin with surrounding single-stranded regions is displayed and region 1 is boxed (nt 237–242). The DIS palindrome (^{257}GCGCGC^{262}) is unreactive in the KL dimer because it is engaged in intermolecular basepairing. The nt color specifies the reactivity (0–0.4: black, 0.4–0.85: orange, and >0.85: red). The mean reactivity values and SD are shown in Supplementary Table 1.

^{278}GG^{279}) or three base pairs (^{240}CUC^{242} interacts with ^{278}GGG^{280}). The 1N and 1P reversions destabilized the mutant DIS hairpin, but only 1N revertants shifted the LDI–BMH equilibrium towards the wt position, either partially (1N7) or completely (1N22) (Table 7). SHAPE-directed RNA structure analysis confirmed the stabilization of the 1N and 1P DIS hairpins and subsequent destabilization in the 1N7 (C–G replaced by less stable U–G base pair) and 1P22 (C–G replaced by A–G mismatch) revertants. Surprisingly, the probing showed a rearranged DIS structure for the 1N22 revertant as the upper and central DIS domains are destabilized compared to 1N (Fig. 4 and Supplementary Table 1). No reactivity changes were observed in other leader regions, indicating that only local DIS changes occur.

The 1N and 1P mutants and revertants produced wt amounts of virus (Fig. 5A). The mutants exhibited strongly delayed replication and all revertants demonstrated improved replication, albeit not as efficient as wt (Fig. 5B). The mutants 1N and 1P exclusively adopt the BMH conformation, unlike LDI folding of wt (Fig. 5C, upper panel). As a consequence, a high level of KL RNA dimers was formed. These effects are partially or completely restored for the 1N revertants 1N7 and 1N22 (Fig. 5C, Table 7). The KL dimer level (1N > 1N7 > 1N22) inversely correlates with replication capacity (1N22 > 1N7 > 1N). The 1P revertant 1P22 does not restore the wt RNA dimerization pattern, which is consistent with the RNAstructure prediction that showed a similar (positive) $\Delta\Delta G_{\mathrm{LDI-BMH}}$ value for the mutant and revertant (Table 7).

The wt, mutant and revertant virions contained a similar level of HIV-1 RNA (Fig. 5D). Analysis of the virion RNA showed a reduced dimer level for all mutant and revertant viruses when compared to wt (Fig. 5E). Thus, although the N revertants demonstrated a partial restoration of the RNA dimerization phenotype in vitro (Fig. 5C), such an effect is not seen in vivo, which likely relates to the modest replication capacity of these revertant viruses. These results demonstrate the importance of a correct RNA secondary structure for in vivo dimerization.

Table 6
Evolution of region 1 mutants.

	Mutant sequence [a]	Replication capacity [b]	Replication capacity upon long-term culture [b]	Days	# cultures	Frequency	First-site mutations [a]	Second-site mutations in leader			Analysis [c]
wt	GCAGGA	++									
1A	UUUGUA	+	+	112	2	1		G348A (Ala5Thr)			
			++	118		1		C152A	C227U		
1B	CUCUUC	––	-	60-136	9	7					
			++	56-60		2					
1C	UAAUUG	-	-	60-90	9	6					
			+	107		1		G275A			*
			+	86		1	UAAUUG → UAAUU**A**	C160A	G224A	G354A (Val7Ile)	***
			+/-	60		1					
1D	AGUAAU	-	+	56-102	9	5					
			-	90		1					
			+	83		1		G121U	C160U		
			+	56		1		U200A			
			+	63		1		insA306			
1E	CUUUGA	+/-	+	90/102	3	2					
			+	118		1		U200C			
1F	AACCCU	-	+/-	103	7	1		G106A			
			+/-	63		3					
			-	60		3					
1G	CAUGCU	-	-	90	8	5					
			+	60-93		2					
			+	60		1		G348A (Ala5Thr)			*
1H	CAUGUA	+/-	+	83-118	3	3					
1I	AACGAA	+/-	+	118	2	1					
			+/-	91		1		G106A	G162A		*
1J	UAUCGA	+/-	++	130	2	2					
1K	CUUGAC	+/-	++	105	3	1		G348A (Ala5Thr)			**
			++	115		1	CUUGAC → **A**UUGAC				**
			++	110		1					
1L	CAAAGA	+/-	++	118	1	1		G369A (Glu12Lys)			
1M	AGCAAC	-	++	102	9	1	AGCAAC → AGCA**U**C	C150A	G213U		*
			+/-	86		1	AGCAAC → AGCA**U**C	ins A305			
			+/-	56-122		6	AGCAAC → AGCA**U**C				*
			+	60		1	AGCAAC → AGCA**U**C	G106A	G164A		
1N	AAUACC	-	++	107	9	1	AAUACC → AAUA**U**C	G106A	C127G/C		**
			++	91		1	AAUACC → AAUA**U/C/G**C	C114A			**
			+	97		3					
			+/-	56		1		C160U			
			-	60		1					

1O	AAGGGA	++	++	105	5	1		G106A	C160C/T	G166A	
			++	83		3					
			+	83		1		G106A	G166A		
1P	UCACUC	-	+	107	9	1		G366A (Gly11Arg)			*
			+	100		1	UCACUC → UCA**A**UC	G106A	G366A (Gly11Arg)		***
			-	60-90		6					
			+/-	56		1					
1Q	UUCCAC	-	-	60-74	9	5					
			+/-	90		1		G369A (Glu12Lys)	C175A		*
			+/-	116		1		G369A (Glu12Lys)			*
			+	60		1		C150U	C175A		*
			+/-	63		1					
1R	CUUCCC	––	-	60-136	9	7					
			+	93		1		G106A	C362A (Ser9Ser)	G369A (Glu12Lys)	*
			+/-	68		1					
1S	ACAGUA	+	+	105	3	1		C349U (Ala5Val)			
			+	83		1	ACAGUA → **G**CAGUA				
			+	122		1	ACAGUA → **G**CAG**G**A				
1T	AGAUAC	+/-	++	107	2	1					
			+/-	129		1		G164A			*
1U	CCCCCA	-	-	60-90	8	6					
			+	60-98		2					
1V	AACAAU	-	+	87	9	1		G164A			*
			+	94		1	AACAAU → **G**ACAAU				***
			+	91		1		C349U (Ala5Val)			*
			+/- or ++	63		5					
			++	56		1	AACAAU → A**C**CAAU				

[a] nt differing from wt are underlined for clarity, revertant mutations in bold.
[b] ++: wt-like, +: slightly delayed compared to wt, +/-: delayed compared to wt, -: no replication in week 1 and 2.
[c] The observed leader mutations were recloned in the original mutant and the replication capacity of the revertant clone was analyzed.
* no improvement.
** improved replication.
*** improved replication, used for detailed analysis.

Table 7
RNA secondary structure and virus replication of designed sequence-only mutants.

Mutant	Sequence[a]	Replication[b]	ΔG_{DIS}[c]	ΔG_{LDI}[c]	ΔG_{BMH}[c]	$\Delta\Delta G_{LDI-BMH}$[d]
wt	GCAGGA	++	−12.9	−139.8	−134.2	−5.6
1N	AAUACC	−	−19.5	−136.4	−141.6	5.2
1N7	AAUA**U**C	+/−	−16.4	−136.4	−136.2	−0.2
1N22	AAUA**G**C	+/−	−16.0	−139.5	−134.0	−5.5
1P	UCACUC	−	−19.3	−136.2	−142.5	6.3
1P22	UCA**A**UC	+/−	−16.4	−134.1	−140.2	6.1
1C	UAAUUG	−	−13.5	−136.3	−137.0	0.7
1C3	UAAUU**A**	+/−	−12.9	−135.9	−135.2	−0.7
1V	AACAAU	−	−13.1	−138.3	−133.8	−4.5
1V7	**G**ACAAU	+/−	−13.1	−138.0	−133.8	−4.2

[a] Nucleotides differing from wt underlined.
[b] Virus replication scored by ++ (day 6/7: CA-p24 >95% of maximum of wt), + (day 6: CA-p24 >50% of maximum of wt), +/− (day 10: CA-p24 >35% of maximum of wt), − (detectable CA-p24 level after day 10).
[c] ΔG in kcal/mol of the DIS hairpin and the complete leader in LDI and BMH conformation as predicted by RNAstructure.
[d] $\Delta\Delta G$ in kcal/mol of the LDI–BMH equilibrium.

In summary, mutants with a stabilized DIS hairpin are hampered in both in vitro and in vivo RNA dimerization and viral replication is blocked. The replication-competent revertants partially restore the hairpin structure and consequently RNA dimerization.

Revertant analysis of region 1 sequence mutants

We also analyzed two mutant-revertant sets for which RNAstructure analysis predicted no or a small effect on DIS hairpin stability and a small or moderate effect on the $\Delta\Delta G_{LDI-BMH}$ equilibrium: mutant 1C (UAAUUG) with revertant 1C3 (UAAUU***A***), and mutant 1V (AACAAU) with revertant 1V7 (***G***ACAAU) (Table 7). These viruses may shed light on the role of the region 1 sequence. For mutant 1C, however, SHAPE-directed structure prediction revealed a rearranged DIS hairpin that no longer exposed the palindromic sequence, while leaving other leader regions unchanged (Fig. 4, Supplementary Table 1). Revertant 1C3 restored the DIS hairpin structure. Mutant 1V and revertant 1V7 extend the DIS hairpin with a single base pair ^{242}U–G^{278} (Fig. 4, Supplementary Table 1), but ^{242}U is more reactive in the revertant, suggesting that the closing base pair ^{242}U–G^{278} is weakened. Other regions of the leader RNA structure were not affected.

Virus production was similar for all constructs (Fig. 6A) and both revertants partially restored virus replication (Fig. 6B). Next we tested the mutant–revertant sets for in vitro RNA dimerization. Mutant 1C RNA adopted the BMH conformation and dimerization was strongly reduced (Fig. 6C). Revertant 1C3 partially restored LDI folding and fully repaired dimer formation. Misfolding of the DIS hairpin in mutant 1C, which is repaired in revertant 1C3, explains these phenotypes (Fig. 4). Mutant 1V formed slightly more BMH than wt and increased the dimer level, which is largely restored in 1V7 (Fig. 6C).

Virion RNA packaging was two-fold reduced for mutant 1C with the rearranged DIS hairpin compared to wt (Fig. 6D). Revertant 1C3 restored RNA packaging levels. The 1V mutant–revertant set did not affect RNA packaging. RNA dimerization was affected for mutants 1C and 1V, as the virions RNA dimer levels were reduced compared to wt (Fig. 6E). Revertant 1C3 does not restore this defect, whereas 1V7 does partially restore the dimer level.

Altogether, despite intensive efforts, 1V7 may be the only structure-neutral mutant left to support a sequence-specific role for region 1 in HIV-1 replication. The 1V7 revertant structure was wt-like (similar ΔG, $\Delta\Delta G_{LDI-BMH}$, SHAPE-directed RNA structure prediction and LDI–BMH equilibrium), but the in vivo RNA dimer level was reduced, perhaps causing the replication defect. This result may suggest a sequence-specific contribution of region 1 in RNA dimer maturation. At the same time, it is becoming increasingly clear that it is very difficult to design strictly structure-neutral mutations in this region of the HIV-1 RNA genome.

Discussion

The non-protein-coding leader is a highly conserved and structured part of the HIV-1 RNA genome that regulates multiple molecular mechanisms during HIV-1 replication. In this study we analyzed in detail the role of three short single-stranded regions flanking the DIS, SD and Ψ hairpins. Region 1 mutations are most detrimental to virus replication, which confirms the results of an in vivo SELEX experiment that documented a particularly strong selection pressure for region 1 (van Bel et al., 2014a). Most region 1 mutants affected the stability of the DIS hairpin and the leader RNA structure, which correlated with poor virus replication. In contrast, most region 2 and 3 mutations did not significantly affect the RNA structure and virus replication. Extreme sequence conservation in region 1 is present among HIV-1 isolates, but considerable sequence conservation is also present in regions 2 and 3 (van Bel et al., 2014a). On the other hand, partial deletions in regions 2 and 3 are observed in a few virus isolates and we scored replication for a partial region 3 deletion mutant (van Bel et al., 2014a). This may suggest that regions 2 and 3 are not absolutely required for HIV-1 replication, at least in vitro. Long-term culturing of region 1 mutants yielded a large set of revertant viruses. Testing of some of these revertants confirmed that a wt-like stability of the DIS hairpin structure is critically important. Mutations that stabilize the DIS hairpin cause a shift in the LDI–BMH equilibrium towards BMH and consequently induce RNA dimerization. DIS hairpin destabilization restores the LDI–BMH equilibrium and RNA dimerization pattern.

This study reinforces the idea that the sequence of region 1 is critical for HIV-1 replication. For instance, mutant viruses frequently reverted towards the wt sequence, consistent with a high selection pressure (van Bel et al., 2014a). We designed additional mutants for which no or a small effect on DIS folding and the leader structure was predicted. RNA dimerization problems were scored for these mutants, suggesting an important sequence contribution. However, RNA structure probing indicated changes in the secondary structure, either subtle nt reactivity changes or base pair rearrangements in the DIS hairpin. The unique mutant that revealed no RNA structure effect is 1V7, although we cannot formally exclude that the mutations induce an effect on the leader RNA conformation. The in vitro ΔG values, RNA dimerization pattern and probed RNA structure of 1V7 are similar to wt, but this mutant is impaired in in vivo dimerization and virus replication. After testing 22 mutants and many revertants, we conclude that it is difficult to change the sequence of region 1 without invoking RNA structural effects. This may also explain the high conservation of region 1 sequences in natural HIV-1 isolates (van Bel et al., 2014a). We previously pointed to the extreme sensitivity of the HIV-1 leader to unwanted RNA structure changes, e.g. TAR mutations frequently induce structural effects in downstream leader domains (Das et al., 2012; Vrolijk et al., 2008).

Long-term culturing of viruses mutated in region 1 or 2 resulted in a wide range of second-site reversions, but most did not significantly improve virus replication. We therefore did not find strong evidence for an interaction between region 1 and upstream/downstream regions. Nevertheless, the reversion pattern demonstrated some specificity. For region 1 mutants, the most frequently observed second-site mutations are G106A (8 ×), G369A (4 ×), C160U (3 ×),

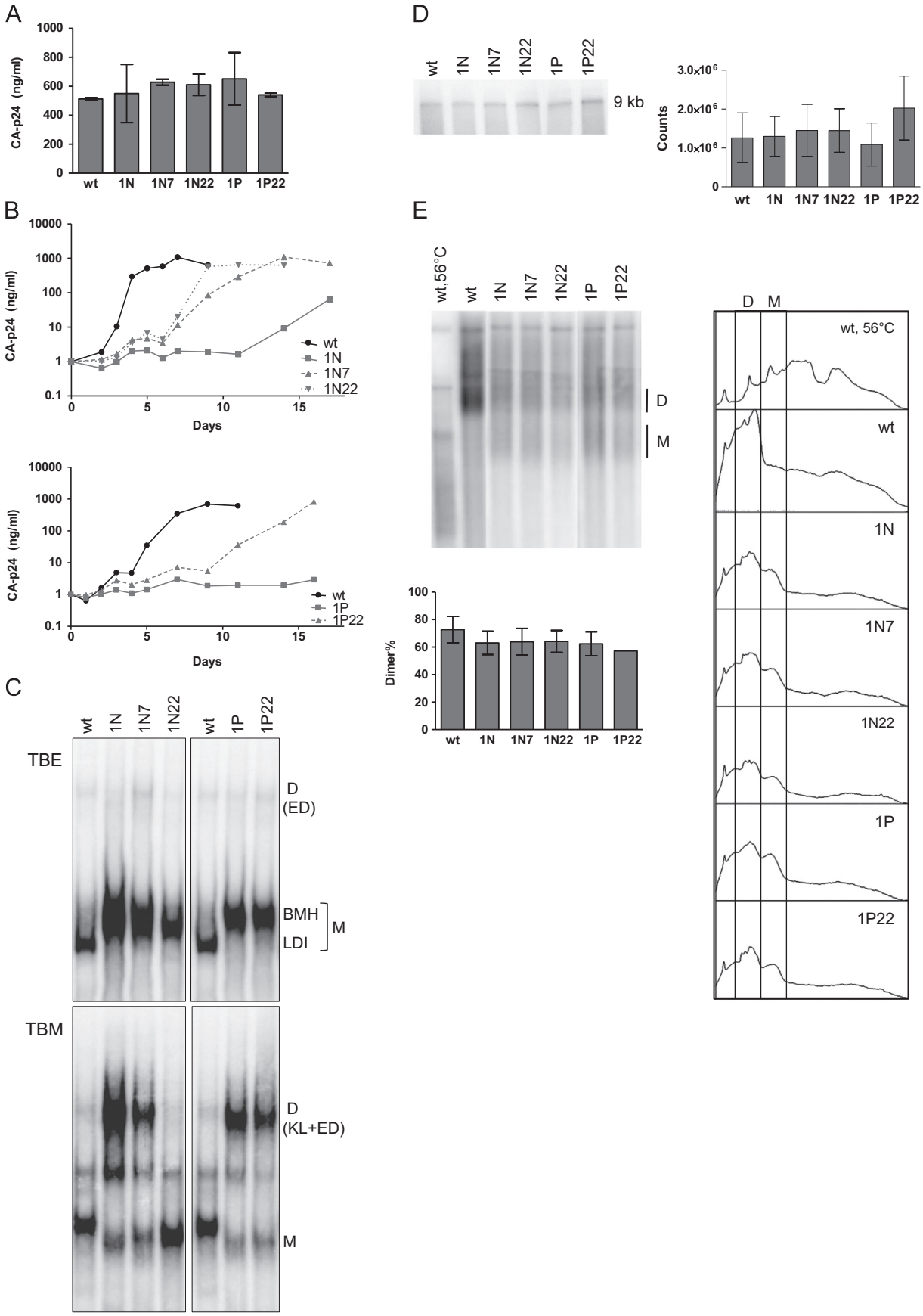

Fig. 5. Characterization of region 1 mutants 1N and 1P (with a stabilized DIS hairpin) and their revertants. For a detailed explanation see the legend to Fig. 3.

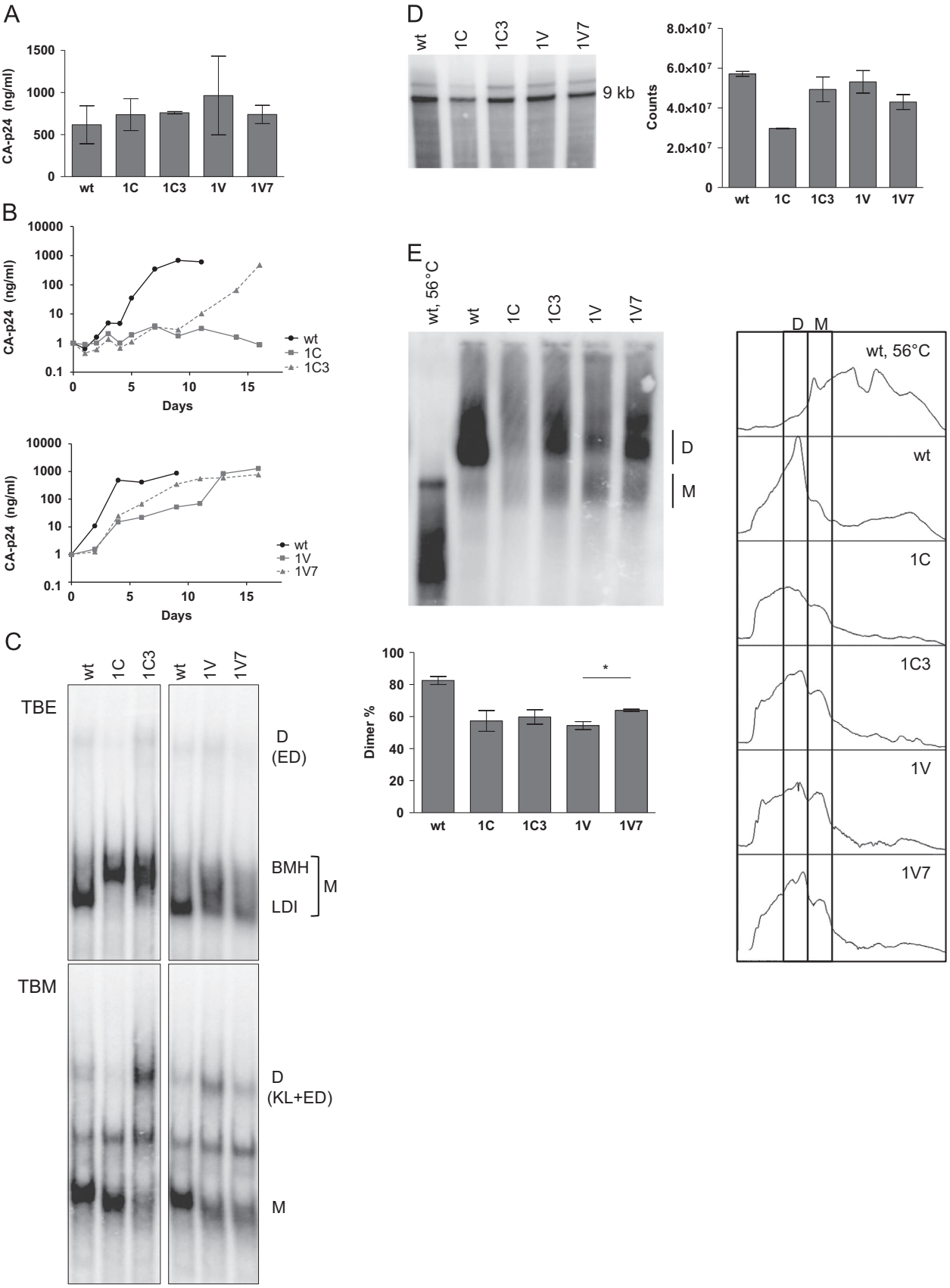

Fig. 6. Characterization of region 1 mutants 1C and 1V (with minimal effects on the DIS hairpin stability) and their revertants. For a detailed explanation see the legend to Fig. 3.

G164A (3 ×) and G348A (3 ×) (Table 6). Interestingly, a similar revertant analysis for region 2 mutants identified other second-site reversions. No obvious positional clustering is apparent, except that two hotspots in Gag are apparent that cause an amino acid substitution: non-silent mutations G348A [A5T] and G369A [E12K]). In addition, the reversions could have RNA structure effects. The frequently observed G106A mutation alters the U5 sequence that was predicted to form a pseudoknot interaction with the DIS palindrome (^{257}GCGCG^{261} interacts with ^{105}UGUGC^{109}) (Lu et al., 2011a). This G106A mutation would likely block DIS-U5 pseudoknot formation. On the other hand, the observed reversions do confirm a predicted pseudoknot interaction between the 5′ end of the gag gene and region 1 (^{234}GACGC^{238} interacts with ^{348}GCGUC^{352}) (Sakuragi et al., 2012). Region 1 mutations can disrupt 2 of the 5 base pairs (^{237}G–C^{352} and ^{238}C–G^{351}). In 4 cultures reversions were selected that partially restored this pseudoknot structure by base pair covariation (1A and 1K: G348A, 1S and 1V: C349U). However, not all evidence supports the pseudoknot relevance as mutant 1O (^{237}AAGGGA^{242}) is not able to form the pseudoknot interaction, yet is replication-competent.

The large set of mutant and revertant viruses tested in this study enables us to start looking for parallels between the in silico, in vitro and in vivo studies. RNA structure prediction and RNA structure probing matched in case of mutants with an extended DIS hairpin of at least 2 base pairs. However, when RNAstructure predicts no or small secondary structure effects, probing data sometimes indicated a more dramatic RNA folding effect. The SHAPE-informed RNA structure predictions were nearly always a predictor of the in vitro RNA dimerization behavior: mutants with a stabilized DIS hairpin fold more BMH, leading to a higher total dimer level. The ED levels seem to depend on the stability of the predicted DIS hairpin, as a too stable DIS hairpin inhibits ED formation, which can be understood as the DIS stem has to open during ED formation. Thus, the DIS hairpin stability is important to maintain the correct levels of monomer (LDI and BMH) and dimer (KL and ED) conformations. In addition, the in vitro dimerization data frequently correlate with the in vivo effects. Mutants with an in vitro RNA dimerization problem are not able to replicate and show an in vivo RNA dimerization defect. Only when the in vitro dimerization pattern is wt-like and the virus replicates efficiently, one can detect high amounts of RNA dimers in virions.

In this study we examined the role of the single-stranded region 1 sequence immediately 5′ of the DIS hairpin in the HIV-1 leader RNA on virus replication. Designing mutants that do not alter the structure seems easy, but proved rather difficult. Sometimes RNA probing was required to reveal such effects, indicating that RNA structure prediction programs do not always provide accurate information. The combined results may cautiously suggest that the wt sequence was uniquely selected from total sequence space during HIV-1 evolution to support the leader RNA structure. We therefore suggest a passive role for the single-stranded region 1 by preventing aberrant folding of the surrounding RNA structures. There is accumulating evidence for such mutation-induced RNA misfolding effects in the HIV-1 leader RNA (Abd El-Wahab et al., 2014; Aldovini and Young, 1990; Clever and Parslow, 1997; Das et al., 2012; Helga-Maria et al., 1999; Lever et al., 1989; Vrolijk et al., 2008). We postulate that such a passive role of a particular RNA sequence is not a HIV-specific property, but applicable to RNA structures in general (Bassi et al., 1999; Levinger et al., 1995; Pan and Woodson, 1999). Most RNA mutagenesis studies neglect to score for adverse RNA secondary structure effects. With recent advances in high-throughput RNA probing, i.e. SHAPE, more RNA structural studies are being performed and this will probably shed more light on the interplay between RNA sequence and structure. In case of HIV-1 RNA, the leader may be particularly sensitive to structural changes because many sequence and structure motifs with essential regulatory functions are clustered in this region. Because the stability of some individual hairpin motifs is delicately balanced (e.g. the polyA, DIS and SD hairpins (Abbink et al., 2005; Clever and Parslow, 1997; Das et al., 1999; Mueller et al., 2014)), the leader RNA seems especially vulnerable to mutation-induced structural effects. We report that even the sequence of small single-stranded leader RNA domains is carefully chosen to avoid unwanted RNA structure effects.

Conclusions

The HIV-1 RNA genome encodes multiple proteins and many RNA signals that control specific steps of the replication cycle. Using a combination of in silico, in vitro methods and in vivo virus experiments, we demonstrate that it is nearly impossible to introduce structure-neutral mutations in the untranslated leader region of the HIV-1 RNA genome. We propose a passive role for the sequence of single-stranded segments to prevent aberrant folding of important structured RNA signals in the surrounding leader RNA.

Materials and methods

Cell culture

HEK 293T cells were grown in DMEM medium (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (FBS), 1X MEM NEAA (Gibco), 40 U/ml penicillin and 40 μg/ml streptomycin at 37 °C and 5% CO_2. SupT1 T cells were grown in advanced RPMI 1640 medium (Gibco) supplemented with 1% FBS, 2 mM L-glutamine (Gibco), 15 U/ml penicillin and 15 μg/ml streptomycin at 37 °C and 5% CO_2.

Virus production, replication and evolution

Plasmids encoding region 1, 2 or 3 mutated HIV-1 variants were randomly picked from previously constructed HIV-1 libraries (based on pLAI (Peden et al., 1991)) with small randomized sequence stretches (van Bel et al., 2014a). Plasmid DNA was isolated using the NucleoSpin® Plasmid kit (Macherey-Nagel) and the leader region was sequenced with primer AD-Gag (Das et al., 1997).

293T cells were cultured to 70% confluency in 2 cm^2 wells and transfected with 1 μg HIV-1 DNA using the calcium-phosphate method (Das et al., 1999). After 16 h, the culture medium was replaced and after an additional 24 h the virus-containing supernatant was harvested. The CA-p24 level was measured by enzyme-linked immunosorbent assay (ELISA) as described (van Bel et al., 2014a). Similar amounts of virus (equivalent to 1 ng CA-p24) were used to infect SupT1 T cells (2×10^5 per ml), which were split in fresh culture medium (1:5) twice a week. Replication was monitored by assessing the CA-p24 level in the culture supernatant. For replication-impaired mutants, we continued the cultures to select revertant viruses that may arise by spontaneous virus evolution (Das and Berkhout, 2010). Upon appearance of virus-induced syncytia, the cell-free culture supernatant was passaged onto fresh SupT1 T cells to passage the virus. This was repeated several times and we eventually isolated the integrated proviral DNA from a massively infected culture. Cells were pelleted and washed once with PBS after which the cells were lysed in 10 mM Tris–HCl pH 8.0, 0.1 mM EDTA, 0.5% Tween20 and 200 μg/ml proteinase K at 56 °C for 60 min. Proteinase K was inactivated by incubation for 10 min at 95 °C. The proviral DNA was PCR amplified with primers U3XbaNotI (Das et al., 2011) and AD-Gag (Das et al., 1997). The

resulting PCR product was sequenced with primer AD-Gag using the BigDye Terminator cycle sequencing kit (Applied Biosystems).

Viral RNA packaging and dimerization

RNA was isolated from HIV-1 virions as described previously (van Bel et al., 2014b). In short, 293T cells were transfected with the pLAI-based constructs using lipofectamine2000 (Invitrogen). The virus-containing supernatant was collected after 48 h and virions were collected over a 20% sucrose cushion by ultracentrifugation in a Beckman SW32 Ti rotor at 32,000 rpm (175,000*g*) for 2 h at 4 °C. The virus pellet was resuspended in lysis buffer (50 mM Tris–HCl pH 7.4, 10 mM EDTA, 100 mM NaCl, 1% SDS and 15 μg GlycoBlue [Ambion]) and a sample was taken for CA-p24 quantitation. Virus was lysed with proteinase K (100 μg/ml) at 37 °C for 30 min and extracted twice with phenol–chloroform–isoamylalcohol (PCI, 25:24:1, Invitrogen) at 4 °C. The RNA was ethanol-precipitated and resolved in 10 mM Tris–HCl pH 7.5, 1 mM EDTA, 100 mM NaCl, 1% SDS, aliquoted and stored at −80 °C.

Northern blot analysis

Denaturing and non-denaturing northern blot analysis was performed as described previously (van Bel et al., 2014c). For the non-denaturing northern blot, equal amounts of viral RNA (equivalent to 250 ng CA-p24) were loaded on a 0.9% agarose gel in 1X TBE and electrophoresed for 6 h at 72 V and 4 °C. The RNA was denatured by soaking the gel in 10% formaldehyde for 30 min at 65 °C. RNA was blotted onto a positively charged nylon membrane (Boehringer Mannheim) using 20X SSC (3.0 M NaCl, 0.3 M sodium citrate, pH 7.0). The RNA was cross-linked to the membrane using an UV crosslinker (Stratagene). The DNA probe consisted of a 1 kbp fragment covering Nef, U3 and R (positions 8770–9784, relative to the transcriptional start site) that was labeled with [α^{32}P]-dCTP (0.33 MBq per μl, Perkin-Elmer) by random-primed labeling (High Prime DNA Labeling Kit, Roche Diagnostics). The membrane was prehybridized and hybridized in ULTRAhyb buffer (Ambion) at 55 °C for 1 h and 16 h, respectively. After extensive washing the blot was analyzed using a phosphorimager (Amersham Biosciences) and the ImageQuant and ImageJ software packages. For the denaturing northern blot, viral RNA (equivalent to 250 ng CA-p24) was mixed with denaturing loading dye (final concentration: 40 mM MOPS pH 7.0, 10 mM sodium acetate, 5% formaldehyde, 0.05 mg/ml ethidium bromide, 0.5 mg/ml orange G, 7 g/ml sucrose) and heated at 65 °C for 10 min before electrophoresis on a 0.9% agarose gel in MOPS buffer with 7% formaldehyde for 4 h at 100 V. Blotting, crosslinking, washing steps and analysis were performed as described for the non-denaturing blot.

In vitro dimerization assay with leader RNA transcripts

LAI-based plasmids with specific leader mutations were used as PCR template with primers A6838 (CACGTCGACTCTAGATAA-TACGACTCACTATA**G**GTCTCTCTGGTTAGACCAGA, T7 RNA polymerase binding site underlined, with the HIV −1 +1 nt transcriptional start site in bold) and MO13 (TCCCCCGCTTAATACTGA). The PCR product encodes the complete HIV-1 leader (nt 1–368) and was purified using the Gel and PCR Clean-up Kit (Macherey-Nagel) followed by ethanol-precipitation. The pellet was dissolved in RNase-free water and in vitro transcribed for 16 h at 37 °C using the MEGAshortscript T7 Kit (Ambion) in the presence of 1 μl [α^{32}P] UTP (0.33 MBq per μl, Perkin-Elmer). The template DNA was degraded by a 30 min incubation at 37 °C with TURBO DNase. The RNA was passaged over a NucAway column (Ambion) to remove unincorporated nucleotides and purified on a denaturing 4% polyacrylamide gel. The 368-nt RNA transcript was eluted from gel in 0.5 M ammonium acetate, 10 mM EDTA and 0.5% SDS for 16 h at 30 °C. The RNA was extracted with PCI (25:24:1), ethanol precipitated and dissolved in RNase-free water. The RNA was quantified by scintillation counting and the quality was tested by analysis of 100 ng RNA on a denaturing 4% polyacrylamide gel. For dimerization analysis, 200 ng RNA in 20 μl dimerization buffer (125 mM KAc, 2.5 mM MgAc, 25 mM HEPES pH 7.0) was incubated at 85 °C for 2 min, at 65 °C for 10 min and then slowly cooled to room temperature in 1 h to allow RNA folding and dimerization. The RNA was mixed with 10 μl non-denaturing sample buffer (30% glycerol, 1X dimerization buffer with 0.025% bromophenol blue dye) and split for loading onto two 4% polyacrylamide gels in either 0.25X TBE (22.5 mM Tris, 22.5 mM boric acid, 0.5 mM EDTA pH 8.0) or 0.25X TBM (22.5 mM Tris, 22.5 mM boric acid, 0.1 mM $MgCl_2$). The gels were run at 150 V for 3 h at room temperature, dried and analyzed using a phosphorimager (Amersham Biosciences) and ImageQuant software.

RNA secondary structure prediction

RNA secondary structure predictions were performed using the RNAstructure program version 5.6 provided by the Mathews lab (http://rna.urmc.rochester.edu/RNAstructure.html) (Reuter and Mathews, 2010). When indicated, SHAPE (Selective 2′-Hydroxyl Acylation analyzed by Primer Extension) reactivity data (see below) was used as pseudo-free energy constraint. Standard settings were used for folding of the DIS hairpin (nt 237–281) and the leader (nt 1–368). When the DIS and polyA hairpin were exposed, the structure was indexed as BMH conformation, otherwise the structure was listed as LDI.

RNA secondary structure probing by SHAPE

SHAPE probing was performed as previously described (Mortimer and Weeks, 2009; Mueller et al., 2014). In short, HIV-1 leader RNA transcripts (nt 1–368) were prepared as described above, but omitting [α^{32}P]UTP. The RNA (3 pmol) was denatured at 90 °C for 3 min in 50 mM HEPES pH 8.0, 200 mM KAc, 5 mM $MgCl_2$, snap-cooled on ice for 1 min and incubated at 37 °C for 1 h to allow folding of the RNA. The RNA was modified with 40 mM BzCN (Sigma-Aldrich) in DMSO or mock-treated with DMSO. RNA was precipitated and reverse transcribed using the FAM-labeled primer AH-AUG368 (TCCCCCGCTTAATACTGACGC) and SuperScript III (Invitrogen). For the sequencing ladder, the RNA template was reverse transcribed with NED-labeled AH-AUG368 primer. The RNA template was degraded by incubation with RNaseH for 20 min at 37 °C and the cDNA was ethanol precipitated. The pellet was dissolved in deionized formamide mixed with GeneScan 500 ROX Size Standard (Applied Biosystems). Samples were analyzed on an Applied Biosystems AB3130 instrument. The QuSHAPE software was used to align the fragments to the reference sequence and the relative peak areas were calculated to determine the reactivity of each nucleotide (Karabiber et al., 2013).

Acknowledgments

We thank Stephan Heynen for performing CA-p24 ELISA. This research was supported by the Netherlands Organization for Scientific Research (NWO-CW, Chemical Sciences Division, Top grant, 700.59.301).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2015.03.050.

References

Abbink, T.E., Berkhout, B., 2003. A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. J. Biol. Chem. 278, 11601–11611.

Abbink, T.E., Ooms, M., Haasnoot, P.C., Berkhout, B., 2005. The HIV-1 leader RNA conformational switch regulates RNA dimerization but does not regulate mRNA translation. Biochemistry 44, 9058–9066.

Abd El-Wahab, E.W., Smyth, R.P., Mailler, E., Bernacchi, S., Vivet-Boudou, V., Hijnen, M., Jossinet, F., Mak, J., Paillart, J.C., Marquet, R., 2014. Specific recognition of the HIV-1 genomic RNA by the Gag precursor. Nat. Commun. 5, 4304.

Aldovini, A., Young, R.A., 1990. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. J. Virol. 64, 1920–1926.

Amarasinghe, G.K., De Guzman, R.N., Turner, R.B., Chancellor, K.J., Wu, Z.R., Summers, M.F., 2000. NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. J. Mol. Biol. 301, 491–511.

Andersen, E.S., Contera, S.A., Knudsen, B., Damgaard, C.K., Besenbacher, F., Kjems, J., 2004. Role of the trans-activation response element in dimerization of HIV-1 RNA. J. Biol. Chem. 279, 22243–22249.

Bassi, G.S., Mollegaard, N.E., Murchie, A.I., Lilley, D.M., 1999. RNA folding and misfolding of the hammerhead ribozyme. Biochemistry 38, 3345–3354.

Berkhout, B., Das, A.T., 2009. Virus evolution as a tool to study HIV-1 biology. Methods Mol. Biol. 485, 436–451.

Berkhout, B., Das, A.T., van Wamel, J.L., 1998. The native structure of the human immunodeficiency virus type 1 RNA genome is required for the first strand transfer of reverse transcription. Virology 249, 211–218.

Berkhout, B., van Wamel, J.L., 1996. Role of the DIS hairpin in replication of human immunodeficiency virus type 1. J. Virol. 70, 6723–6732.

Berkhout, B., van Wamel, J.L., 2000. The leader of the HIV-1 RNA genome forms a compactly folded tertiary structure. RNA 6, 282–295.

Chen, J., Nikolaitchik, O., Singh, J., Wright, A., Bencsics, C.E., Coffin, J.M., Ni, N., Lockett, S., Pathak, V.K., Hu, W.S., 2009. High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. Proc. Natl. Acad. Sci. USA 106, 13535–13540.

Chin, M.P., Rhodes, T.D., Chen, J., Fu, W., Hu, W.S., 2005. Identification of a major restriction in HIV-1 intersubtype recombination. Proc. Natl. Acad. Sci. USA 102, 9002–9007.

Clever, J.L., Parslow, T.G., 1997. Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. J. Virol. 71, 3407–3414.

Darlix, J.L., de Rocquigny, H., Mauffret, O., Mely, Y., 2014. Retrospective on the all-in-one retroviral nucleocapsid protein. Virus Res. 193, 2–15.

Das, A.T., Berkhout, B., 2010. HIV-1 evolution: frustrating therapies, but disclosing molecular mechanisms. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 365, 1965–1973.

Das, A.T., Harwig, A., Berkhout, B., 2011. The HIV-1 Tat protein has a versatile role in activating viral transcription. J. Virol. 85, 9506–9516.

Das, A.T., Klaver, B., Berkhout, B., 1999. A hairpin structure in the R region of the human immunodeficiency virus type 1 RNA genome is instrumental in polyadenylation site selection. J. Virol. 73, 81–91.

Das, A.T., Klaver, B., Klasens, B.I., van Wamel, J.L., Berkhout, B., 1997. A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. J. Virol. 71, 2346–2356.

Das, A.T., van Dam, A.P., Klaver, B., Berkhout, B., 1998. Improved envelope function selected by long-term cultivation of a translation-impaired HIV-1 mutant. Virology 244, 552–562.

Das, A.T., Vrolijk, M.M., Harwig, A., Berkhout, B., 2012. Opening of the TAR hairpin in the HIV-1 genome causes aberrant RNA dimerization and packaging. Retrovirology 9, 59.

Didierlaurent, L., Racine, P.J., Houzet, L., Chamontin, C., Berkhout, B., Mougel, M., 2011. Role of HIV-1 RNA and protein determinants for the selective packaging of spliced and unspliced viral RNA and host U6 and 7SL RNA in virus particles. Nucleic Acids Res. 39, 8915–8927.

Foley, B., Leitner, T., Apetrei, C., Hahn, B.H., Mizrachi, I., Mullins, J., Rambaut, A., Wolinsky, S., Korber, B., 2013. HIV Sequence Compendium 2013. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, NM, LA-UR, pp. 13–26007.

Fu, W., Gorelick, R.J., Rein, A., 1994. Characterization of human immunodeficiency virus type 1 dimeric RNA from wild-type and protease-defective virions. J. Virol. 68, 5013–5018.

Hache, G., Abbink, T.E., Berkhout, B., Harris, R.S., 2009. Optimal translation initiation enables Vif-deficient human immunodeficiency virus type 1 to escape restriction by APOBEC3G. J. Virol. 83, 5956–5960.

Haddrick, M., Lear, A.L., Cann, A.J., Heaphy, S., 1996. Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. J. Mol. Biol. 259, 58–68.

Helga-Maria, C., Hammarskjold, M.L., Rekosh, D., 1999. An intact TAR element and cytoplasmic localization are necessary for efficient packaging of human immunodeficiency virus type 1 genomic RNA. J. Virol. 73, 4127–4135.

Heng, X., Kharytonchyk, S., Garcia, E.L., Lu, K., Divakaruni, S.S., LaCotti, C., Edme, K., Telesnitsky, A., Summers, M.F., 2012. Identification of a minimal region of the HIV-1 5′-leader required for RNA dimerization, NC binding, and packaging. J. Mol. Biol. 417, 224–239.

Hill, M.K., Shehu-Xhilaga, M., Campbell, S.M., Poumbourios, P., Crowe, S.M., Mak, J., 2003. The dimer initiation sequence stem-loop of human immunodeficiency virus type 1 is dispensable for viral replication in peripheral blood mononuclear cells. J. Virol. 77, 8329–8335.

Hoglund, S., Ohagen, A., Goncalves, J., Panganiban, A.T., Gabuzda, D., 1997. Ultrastructure of HIV-1 genomic RNA. Virology 233, 271–279.

Houzet, L., Paillart, J.C., Smagulova, F., Maurel, S., Morichaud, Z., Marquet, R., Mougel, M., 2007. HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. Nucleic Acids Res. 35, 2695–2704.

Huthoff, H., Berkhout, B., 2001a. Mutations in the TAR hairpin affect the equilibrium between alternative conformations of the HIV-1 leader RNA. Nucleic Acids Res. 29, 2594–2600.

Huthoff, H., Berkhout, B., 2001b. Two alternating structures of the HIV-1 leader RNA. RNA 7, 143–157.

Huthoff, H., Berkhout, B., 2002. Multiple secondary structure rearrangements during HIV-1 RNA dimerization. Biochemistry 41, 10439–10445.

Jalalirad, M., Saadatmand, J., Laughrea, M., 2012. Dominant role of the 5′ TAR bulge in dimerization of HIV-1 genomic RNA, but no evidence of TAR–TAR kissing during in vivo virus assembly. Biochemistry 51, 3744–3758.

Jones, K.L., Sonza, S., Mak, J., 2008. Primary T-lymphocytes rescue the replication of HIV-1 DIS RNA mutants in part by facilitating reverse transcription. Nucleic Acids Res. 36, 1578–1588.

Jouvenet, N., Simon, S.M., Bieniasz, P.D., 2009. Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. Proc. Natl. Acad. Sci. USA 106, 19114–19119.

Karabiber, F., McGinnis, J.L., Favorov, O.V., Weeks, K.M., 2013. QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. RNA 19, 63–73.

Kasprzak, W., Bindewald, E., Shapiro, B.A., 2005. Structural polymorphism of the HIV-1 leader region explored by computational methods. Nucleic Acids Res. 33, 7151–7163.

Kozak, M., 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. 266, 19867–19870.

Laughrea, M., Jette, L., 1996. Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNAs can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248–271 are dispensable for dimer formation. Biochemistry 35, 1589–1598.

Lever, A., Gottlinger, H., Haseltine, W., Sodroski, J., 1989. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J. Virol. 63, 4085–4087.

Levin, J.G., Guo, J., Rouzina, I., Musier-Forsyth, K., 2005. Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism. Prog. Nucl. Acid Res. Mol. Biol. 80, 217–286.

Levinger, L., Vasisht, V., Greene, V., Bourne, R., Birk, A., Kolla, S., 1995. Sequence and structure requirements for Drosophila tRNA 5′- and 3′-end processing. J. Biol. Chem. 270, 18903–18909.

Liang, C., Rong, L., Cherry, E., Kleiman, L., Laughrea, M., Wainberg, M.A., 1999a. Deletion mutagenesis within the dimerization initiation site of human immunodeficiency virus type 1 results in delayed processing of the p2 peptide from precursor proteins. J. Virol. 73, 6147–6151.

Liang, C., Rong, L., Laughrea, M., Kleiman, L., Wainberg, M.A., 1998. Compensatory point mutations in the human immunodeficiency virus type 1 Gag region that are distal from deletion mutations in the dimerization initiation site can restore viral replication. J. Virol. 72, 6629–6636.

Liang, C., Rong, L., Quan, Y., Laughrea, M., Kleiman, L., Wainberg, M.A., 1999b. Mutations within four distinct gag proteins are required to restore replication of human immunodeficiency virus type 1 after deletion mutagenesis within the dimerization initiation site. J. Virol. 73, 7014–7020.

Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E.L., Kharytonchyk, S., Dorjsuren, B., Kulandaivel, G., Jones, S., Hiremath, A., Divakaruni, S.S., LaCotti, C., Barton, S., Tummillo, D., Hosic, A., Edme, K., Albrecht, S., Telesnitsky, A., Summers, M.F., 2011a. NMR detection of structures in the HIV-1 5′-leader RNA that regulate genome packaging. Science 334, 242–245.

Lu, K., Heng, X., Summers, M.F., 2011b. Structural determinants and mechanism of HIV-1 genome packaging. J. Mol. Biol. 410, 609–633.

Moore, M.D., Fu, W., Nikolaitchik, O., Chen, J., Ptak, R.G., Hu, W.S., 2007. Dimer initiation signal of human immunodeficiency virus type 1: its role in partner selection during RNA copackaging and its effects on recombination. J. Virol. 81, 4002–4011.

Moore, M.D., Hu, W.S., 2009. HIV-1 RNA dimerization: it takes two to tango. AIDS Rev. 11, 91–102.

Mortimer, S.A., Weeks, K.M., 2009. Time-resolved RNA SHAPE chemistry: quantitative RNA structure analysis in one-second snapshots and at single-nucleotide resolution. Nat. Protoc. 4, 1413–1421.

Mueller, N., van Bel, N., Berkhout, B., Das, A.T., 2014. HIV-1 splicing at the major splice donor site is restricted by RNA structure. Virology 468–470, 609–620.

Muriaux, D., De, R.H., Roques, B.P., Paoletti, J., 1996a. NCp7 activates HIV-1Lai RNA dimerization by converting a transient loop–loop complex into a stable dimer. J. Biol. Chem. 271, 33686–33692.

Muriaux, D., Fosse, P., Paoletti, J., 1996b. A kissing complex together with a stable dimer is involved in the HIV-1Lai RNA dimerization process in vitro. Biochemistry 35, 5075–5082.

Nikolaitchik, O.A., Dilley, K.A., Fu, W., Gorelick, R.J., Tai, S.H., Soheilian, F., Ptak, R.G., Nagashima, K., Pathak, V.K., Hu, W.S., 2013. Dimeric RNA recognition regulates HIV-1 genome packaging. PLoS Pathog. 9, e1003249.

Ohishi, M., Nakano, T., Sakuragi, S., Shioda, T., Sano, K., Sakuragi, J., 2011. The relationship between HIV-1 genome RNA dimerization, virion maturation and infectivity. Nucleic Acids Res. 39, 3404–3417.
Ooms, M., Huthoff, H., Russell, R., Liang, C., Berkhout, B., 2004. A riboswitch regulates RNA dimerization and packaging in human immunodeficiency virus type 1 virions. J. Virol. 78, 10814–10819.
Paillart, J.C., Shehu-Xhilaga, M., Marquet, R., Mak, J., 2004. Dimerization of retroviral RNA genomes: an inseparable pair. Nat. Rev. Microbiol. 2, 461–472.
Pan, J., Woodson, S.A., 1999. The effect of long-range loop-loop interactions on folding of the Tetrahymena self-splicing RNA. J Mol Biol 294, 955–965.
Peden, K., Emerman, M., Montagnier, L., 1991. Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1LAI, HIV-1MAL, and HIV-1ELI. Virology 185, 661–672.
Reuter, J.S., Mathews, D.H., 2010. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinf. 11, 129.
Sakuragi, J., Shioda, T., Panganiban, A.T., 2001. Duplication of the primary encapsidation and dimer linkage region of human immunodeficiency virus type 1 RNA results in the appearance of monomeric RNA in virions. J. Virol. 75, 2557–2565.
Sakuragi, J., Ueda, S., Iwamoto, A., Shioda, T., 2003. Possible role of dimerization in human immunodeficiency virus type 1 genome RNA packaging. J. Virol. 77, 4060–4069.
Sakuragi, J.I., Ode, H., Sakuragi, S., Shioda, T., Sato, H., 2012. A proposal for a new HIV-1 DLS structural model. Nucleic Acids Res. 40, 5012–5022.
Shen, N., Jette, L., Wainberg, M.A., Laughrea, M., 2001. Role of stem B, loop B, and nucleotides next to the primer binding site and the kissing-loop domain in human immunodeficiency virus type 1 replication and genomic-RNA dimerization. J. Virol. 75, 10543–10549.
Skripkin, E., Paillart, J.C., Marquet, R., Ehresmann, B., Ehresmann, C., 1994. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. Proc. Natl. Acad. Sci. USA 91, 4945–4949.
Song, R., Kafaie, J., Laughrea, M., 2008. Role of the 5′ TAR stem-loop and the U5–AUG duplex in dimerization of HIV-1 genomic RNA. Biochemistry 47, 3283–3293.
Song, R., Kafaie, J., Yang, L., Laughrea, M., 2007. HIV-1 viral RNA is selected in the form of monomers that dimerize in a three-step protease-dependent process; the DIS of stem-loop 1 initiates viral RNA dimerization. J. Mol. Biol. 371, 1084–1098.
van Bel, N., Das, A.T., Berkhout, B., 2014a. In vivo SELEX of single-stranded domains in the HIV-1 leader RNA. J. Virol. 88, 1870–1880.
van Bel, N., Das, A.T., Cornelissen, M.T., Abbink, T.E., Berkhout, B., 2014b. A short sequence motif in the 5′ leader of the HIV-1 genome modulates extended RNA dimer formation and virus replication. J. Biol. Chem. 289, 35061–35074.
van Bel, N., van der Velden, Y., Bonnard, D., Le Rouzic, E., Das, A.T., Benarous, R., Berkhout, B., 2014c. The allosteric HIV-1 integrase inhibitor BI-D affects virion maturation but does not influence packaging of a functional RNA genome. PLoS One 9, e103552.
Vrolijk, M.M., Ooms, M., Harwig, A., Das, A.T., Berkhout, B., 2008. Destabilization of the TAR hairpin affects the structure and function of the HIV-1 leader RNA. Nucleic Acids Res. 36, 4352–4363.
Wilkinson, K.A., Gorelick, R.J., Vasa, S.M., Guex, N., Rein, A., Mathews, D.H., Giddings, M.C., Weeks, K.M., 2008. High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. PLoS Biol. 6, e96.